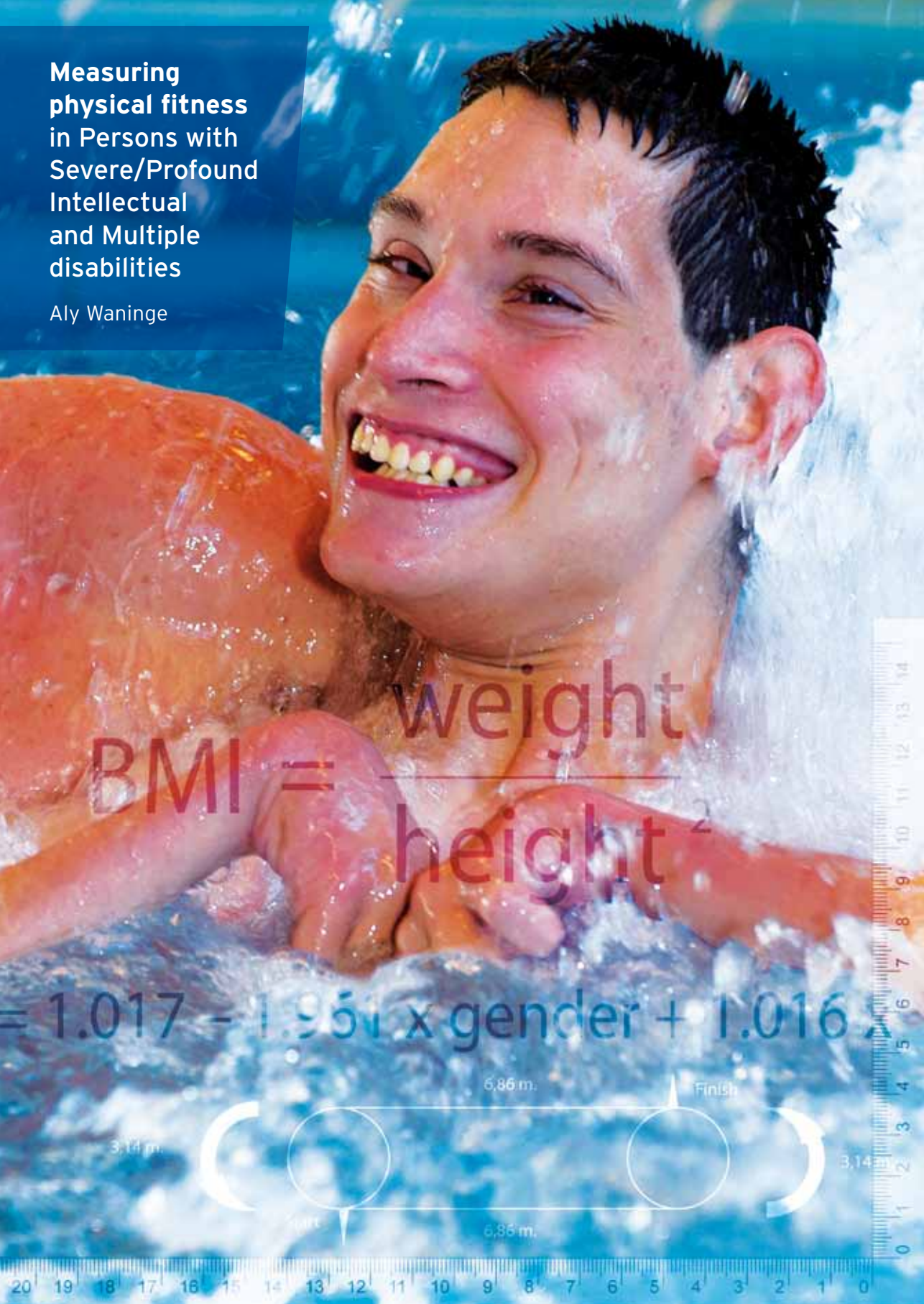
Measuring physical fitness in Persons with Severe/Profound Intellectual and Multiple disabilities

Aly Waninge



Measuring physical fitness
in persons with severe or profound intellectual and multiple disabilities

Aly Waninge

The study presented in this thesis was performed at the Royal Dutch Visio De Brink and at the Research and Innovation Group in Health Care and Nursing of Hanze University of Applied Sciences Groningen, the Netherlands

Aly Waninge

Measuring physical fitness in persons with severe or profound intellectual and multiple disabilities

Dissertation University of Groningen, May 2011

ISBN 978-90-367-4881-0

Graphic Design: Jodh | www.jodh.nl

Photographer: K. Dallinga
Photo cover: Joey Otter

Press: Drukkerij Wilco, Amersfoort

Rijksuniversiteit Groningen

Measuring physical fitness
in persons with severe or profound intellectual and multiple disabilities

Proefschrift

ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
maandag 9 mei 2011
om 16.15 uur.

door

Aly Waninge
geboren op 25 juni 1964
te Zuidlaren

Promotores

Prof. Dr C.P. van der Schans
Prof. Dr B. Steenbergen

Copromotor

Dr R. van Wijck

Beoordelingscommissie

Prof. Dr H.M. Evenhuis
Prof. Dr C. Vlaskamp
Prof. Dr G.E. Lancioni

Paranimfen

Gineke Hanzen
Theresa Berg

Acknowledgements

This study was financially supported by:
- Royal Dutch Visio De Brink in Vries
- Hanze University of Applied Sciences Groningen
- Regional Action-and-Attention Knowledge Circulation (RAAK)
- Fonds Nuts-Ohra, Zoetermeer

Additional financial support for the printing of this thesis has been kindly provided by:
- Barth Fidder/Shapemaster Benelux
- Shapemaster United Kingdom
- Landelijke Stichting voor Blinden en Slechtzienden, Utrecht
- RijksUniversiteit Groningen, Universitair Medisch Centrum Groningen, Graduate School of Medical Sciences

Contents

List of abbreviations

ID	Intellectual Disabilities
SPIMD	Severe or Profound Intellectual and Visual (Multiple) Disabilities
SIMD	Severe Intellectual and Visual (Multiple) Disabilities
PIMD	Profound Intellectual and Visual (Multiple) Disabilities
WHO	World Health Organisation
GMFCS	Gross Motor Function Classification System
ACSM	American College of Sports Medicine
BMI	Body Mass Index
6MWD	6 Minute Walking Distance
aSRT	adapted Shuttle Run Test
ISWT	Incremental Speed Walking Tests
GXT	Graded Treadmill Exercise Test
HR_{peak}	Peak Heart Rate
SMBT	Supra Maximal Block Test
CP	Cerebral Palsy
BBS	Berg Balance Scale
mBBS	modified Berg Balance Scale
TUG	Timed Up and Go Test
POMA	Performance Oriented Mobility Assessment
FICSIT-4	Frailty and Injuries: Cooperative Studies of Intervention Techniques
MAS	Modified Ashworth Scale
MTS	Modified Tardieu Scale

Chapter 1

Introduction

Outline of the introduction

The introduction describes the characteristics of persons with severe or profound intellectual disabilities (ID), as well as the consequences in functioning of additional visual impairments. Firstly, this introduction exposes the various health threats associated with severe or profound ID and/or visual impairments. Next, an examination of the theoretical framework of this thesis is put forward. Finally, both the research questions and the outline of the thesis will be briefly described.

Intellectual disabilities

Persons with ID have significant limitations in both intellectual functioning and adaptive behaviour as expressed in conceptual, social, and practical skills [1]. Intellectual disability is a condition that affects people's ability to make self-determined choices. In addition, people with intellectual disabilities are in danger of being excluded from many situations and opportunities usually available to people not suffering from ID [1]. Based on the WHO population prevalence estimate, the prevalence of ID in the population of Europe is about 1% [2].

The ICD-10 (World Health Organization, WHO) [3] distinguishes four levels of ID: mild (IQ 50-69), moderate (IQ 35-49), severe (IQ 20-34) or profound (IQ under 20). Adults with severe ID have an intellectual age from 3 to 6 years, which is likely to result in a continuous need for support. Adults with profound ID have an intellectual age below 3 years, which results in serious limitations in self-care, continence, communication and mobility [3].

Intellectual disabilities and visual impairment

In all subgroups with ID, prevalence of visual impairment and blindness are significantly higher, compared to the general Dutch population [4]. The severity of the visual impairment is related to the severity of ID. Moreover, prevalence of visual impairments in persons with severe or profound ID is 92% [5]. As the combination of ID and visual impairment is even more detrimental, thereby creating less opportunity for compensation [6], the combination of visual impairment with ID aggravates problems in daily functioning [7].

Health threats associated with ID and/or visual impairments

Research has shown individuals with ID to have twice as much health problems and significantly higher levels of co-morbidity when compared to the general population [8]. As an example, prevalence of neurological problems in persons with ID is 15%, versus 5% in the general population [8]. Moreover, 75% of the persons with severe or profound ID also suffer from locomotor disabilities [9], while adults with mild or moderate ID score significantly lower than a control group without ID on all sensorimotor tests [10]. Also, Shinkfield et al. [11] reported that individuals with mild or moderate ID display inadequacies both in perception as in motor-reproduction.

In addition, those classified with ID are more prone to experience lifestyle related diseases such as diabetes mellitus II or cardiovascular diseases [8, 12, 13]. These persons often suffer from overweight [14, 15, 16] or malnutrition [17]. Obesity in women and underweight in both men and women are also known to be more common in adults with ID than in the general population [12]. Furthermore, Mc Guire et al. [13] found that 68% of their ID sample was overweight or obese. In the Netherlands, over 40% of adults with an intellectual disability were shown to have

overweight [15]. This prevalence is similar to that in other countries [14, 16]. Moreover, persons with ID are often not sufficiently active to achieve health benefits [14, 16, 18, 19], and more than 50 % of the persons with ID of all age categories in Europe have a sedentary lifestyle [20]. As a consequence, these persons may have poor physical fitness [14, 16].

Similar to individuals with ID, persons with visual impairments display poor performance on locomotor skills [21] and have low levels of habitual activity [22], resulting in poor physical fitness when compared to the control group, in this case persons with normal eyesight [23, 24]. Furthermore, persons displaying a combination of severe or profound intellectual disability and a visual impairment are particularly at risk to develop deficits in both locomotor skills as in daily functioning [7]. The combination of these findings puts forward the suggestion that persons having severe or profound intellectual and visual disabilities are likely to display insufficient physical fitness

Terminology relating to persons with severe or profound ID

A wide range of terms is being used to describe persons having a combination of severe or profound intellectual and additional disabilities. The persons studied in this thesis have severe or profound intellectual as well as visual disabilities. In general, the study population is referred to as persons with severe or profound intellectual and multiple disabilities (SPIMD). In the studies examining a population consisting in majority of persons with severe intellectual disabilities, the term severe intellectual and multiple disabilities (SIMD) is used. In the studies examining a population consisting in majority of persons with profound intellectual disabilities, the term profound intellectual and multiple disabilities (PIMD) is used. The term 'multiple' indicates locomotor disabilities, neurological problems, sensory disabilities, and/or problems with food ingestion.

As locomotor skills may influence protocols for measuring physical fitness, it is useful to classify persons with severe or profound ID according to their locomotor skills. The Gross Motor Function Classification System (GMFCS) [25] is a five-level system used to classify the locomotor skills of people with physical disabilities and is also applicable for persons with ID. Participants with a "Level I" classification can generally walk without restrictions but tend to have limitations in some more advanced motor skills. Participants with a "Level II" classification can walk with slight restrictions and do not spontaneously increase their speed during walking. Participants with a "Level III" are only able to walk with walking devices and have restrictions in walking outside as well as in their living environment. Participants with a "Level IV" have limited mobility, but might be able to stand during transfers. Usually they use a wheelchair, which they may drive themselves by hand or by assistive technology. Participants with a "Level V" classification generally have very limited mobility, even with the use of assistive technology. These participants always use a wheelchair.

It is often assumed that persons with profound ID automatically have low locomotor levels and are not able to walk. However, the ability to walk varies considerably in persons with severe ID as well as in persons with profound ID. For example, 75 percent of persons with severe ID is able to walk at least with walking devices (GMFCS I-III), whereas 25 percent is not able to walk (GMFCS IV-V). Moreover, 56 percent of persons with profound ID is able to walk at least with walking devices (GMFCS I-III), whereas 44 percent is not able to walk (GMFCS IV-V). Thus, contrary to common beliefs, it is necessary to perform research in persons with severe or profound ID yet ranging in GMFCS levels from I to V.

Physical fitness and persons with both severe or profound ID and visual impairment

As a sufficient physical fitness level and physical activity improve health [26], and sufficient health in turn improves well-being and quality of life [27, 28, 29], it is imperative to gain comprehensive insight into the physical fitness of persons with SPIMD.
However, the feasibility and reliability of physical fitness measurements and tests in participants with SPIMD have until now not been properly scrutinized, resulting in little reliable knowledge on the physical fitness levels and locomotor skills of persons with SPIMD.

Due to limitations both in intellectual functioning as in adaptive behaviour related to SPIMD, the level of health-related physical fitness is difficult to quantify in a feasible and reliable manner [1]. Therefore, improving feasibility of physical fitness tests in persons with SPIMD needs to be prioritized. Persons with SPIMD are not accustomed to the assessments, have difficulty comprehending what is required of them [30] and often cannot understand instructions [3]. Furthermore, persons with visual disabilities cannot see how test tasks need to be performed [4], hence showing them how to perform the task at hand is useless. In general, if a participant does not understand the tasks within a certain test, the test will automatically fail to provide a realistic impression of the capabilities of the participant, rendering the test invalid. Thus, test instructions for persons with SPIMD require our special focus.

Other factors of influence when determining the feasibility, reliability and validity of physical fitness tests in persons with SPIMD are the prevalence of locomotor disabilities and motivational problems. Adapted test procedures and specific inclusion criteria are required because persons with intellectual, visual, and locomotor disabilities are not able to stand straight or to stand at all [31]. Also, persons with SPIMD are often not motivated to exert themselves fully, which necessitates adjustments to and familiarization with test protocols.

Since physical fitness is related to physical activity [26], it is important to gain insight into the physical activity level in persons with SPIMD. However, as almost 40 % of SPIMD population is simply not able to walk, walking fails to be an adequate representation of a person's overall activity level [32].

Moreover, the presumed low levels of activity in persons with such profound disabilities are often not accurately presented by measurement devices, like activity monitors, which are relatively insensitive [32]. Heart rate monitoring may be an indicator of activity levels assuming a relationship between activity intensity and heart rate [33, 34]. Heart rate monitoring appears to be sufficiently valid for creating broad physical activity categories (e.g. highly active, somewhat active, sedentary) [35]. However, a proper method for dating heart rate patterns in persons with PIMD, as well as the proper correlation between heart rate monitoring and activity levels for this specific group have so far not been examined.

Theoretical framework of the study

International Classification of Functioning, Disability and Health

Physical fitness is related to health [26] which in turn is related to participation [36]. Since participation for persons with ID is important, it is necessary to describe the relation between health and participation for this specific group as well [6].

The International Classification of Functioning, Disability and Health [36] is a commonly used model for various target groups in the field of health care. The concept of participation is defined within the framework of the ICF [3]. Kiestra [6] described participation in persons with a profound intellectual and visual disability as the extent to which someone can take part in or has influence on situations and contexts that are important to him or her, or are considered to be important to him or her by his or her representative or personal coach. This includes situations such as living habits, daily activities, leisure activities, recreation, sports, etc. The level of participation is linked to the abilities of performing the activities in question. In figure 1 the physical fitness components and their related activities are integrated into the model of the ICF to show their relation.

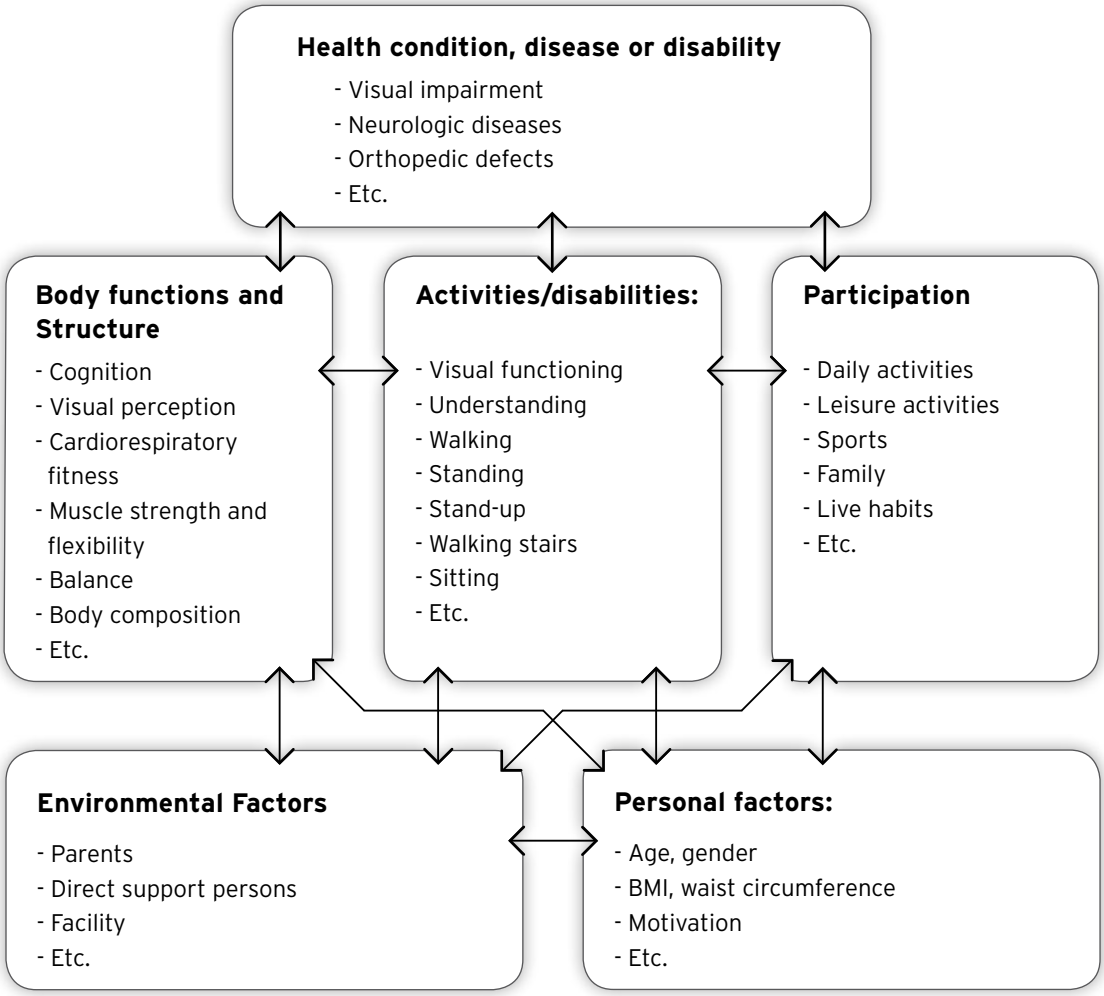


Figure 1. ICF model with physical fitness components and their related activities.

Health, physical fitness, physical activity and quality of life

Several models and concepts have been developed to describe quality of life, participation, physical well-being, physical fitness, physical activity, health and their mutual relatedness. To illustrate the connections between these concepts, a combination of three models is made, which is shown in Figure 2.

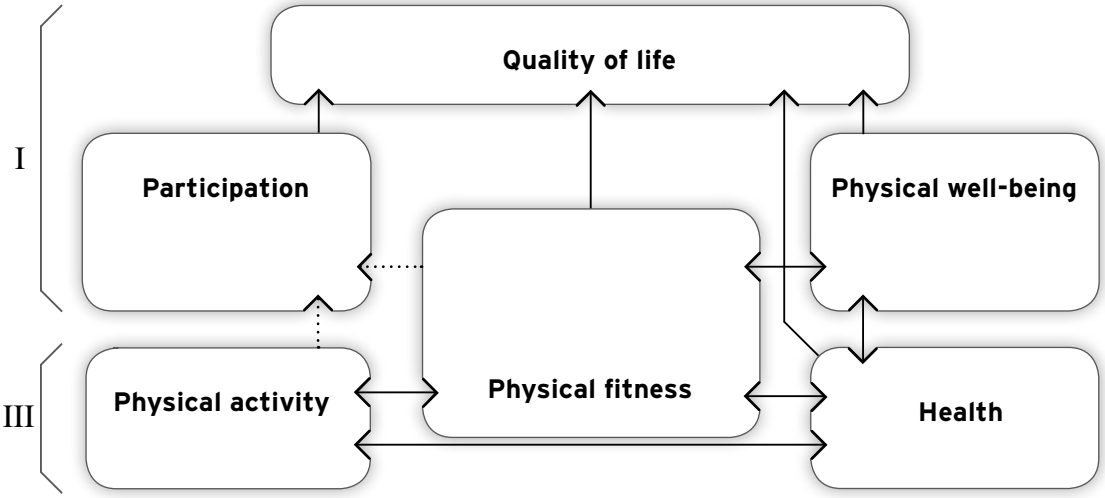


Figure 2. Integration of models and concepts of participation, quality of life, physical well-being, physical activity, physical fitness, and health [1, 26, 27, 28, 29]. I. refers to the description of the model of Shalock in the text; and III refers to the model of Bouchard in the text.

I. First, Schalock et al. [1] provided a concept of quality of life within the international field of intellectual disabilities. Most quality of life concepts share the following common features: general feelings of well-being, feelings of positive social involvement, and opportunities to achieve personal potential [1]. In this model, physical well-being is incorporated as one of the seven domains that contribute to quality of life. Furthermore, quality of life is considered to be an outcome measure of participation [36].

II. Second, we incorporated the following statement into the model: sufficient health improves well-being and quality of life as well [27, 28, 29].

III. The third part of our model reflects the Toronto model, which describes the relation between physical fitness, physical activity and health. Both sufficient physical fitness and physical activity improve health [26]. Physical activity is defined as any body movement produced by skeletal muscles that results in energy expenditure [37], while physical fitness is defined as the ability to perform physical activity, depending on a specific set of attributes that people have or achieve (The American College of Sports Medicine, ACSM) [38]. However, only physical activity which reveals heart rates of more than 55% of the heart rate reserve during 5 days in a week, may gain profit for physical fitness [39]. Health is defined as a state of complete physical, mental and social well-being, and is a positive concept emphasizing social and personal resources, as well as physical capacities [40]. As far as we know, the direct relation between physical activity / physical fitness and participation in persons with SPIMD is still unknown, as indicated by the dotted line.

Components of physical fitness

The attributes of physical fitness can be defined differently for different target groups [U.S. Centers for Disease Control and Prevention] and therefore, physical fitness for persons with SPIMD needs to be described. Hilgenkamp et al. [42] stated that "physical fitness describes how "fit" a person physically is to cope with the demands set by his or her environment" and described physical fitness for older people with ID in a model (table 1) [U.S. Centers for Disease Control and Prevention; 26, 41, 42].

Based on this model, the required attributes of physical fitness for persons with SPIMD are described by caregivers, professionals and scientists in the field of SPIMD. Coordination, reaction time and muscle endurance are considered irrelevant attributes for individuals with such limited cognitive and physical skills.

Caregivers of persons with profound intellectual, visual and locomotor disabilities (profound intellectual and multiple disabilities, PIMD) often describe the quality of daily movements in terms of 'flexibility' or 'stiffness'. Since muscular flexibility is one of the defined physical fitness components for persons with PIMD, muscle tonus or level of spasticity may be used as outcome measures to objectify the concepts of 'flexibility' and 'stiffness'.

Hence, the required attributes of physical fitness for persons with SPIMD are body composition, cardiorespiratory fitness, balance, muscle strength and muscle flexibility (table 1).

Table 1. Model of components of physical fitness for older persons with ID [U.S. Centers for Disease Control and Prevention; 26, 41, 42].

Health-related physical fitness Bouchard et al. (1994)	**U.S. Centers for Disease Control and Prevention**	**Physical fitness of older adults with ID (Hilgenkamp et al. 2010)**	**Physical fitness of persons with PIMD**
Motor	Coordination Reaction time Balance	Coordination Reaction time Balance	 Balance
Muscular	Muscular strength Muscular endurance Flexibility	Muscular strength Muscular endurance Flexibility	Muscular strength Flexibility
Cardiorespiratory	Cardiorespiratory fitness	Cardiorespiratory fitness	Cardiorespiratory fitness
Morphological Metabolic	Body composition		Body composition

Aims and research questions of this thesis

Until present, the feasibility and reliability of physical fitness measurements and tests for participants with SPIMD are unknown. Consequently, knowledge of the physical fitness levels and locomotor skills of persons with SPIMD is scarce. Yet, only with feasible and reliable tests the evaluation of a specific training intervention aimed at promoting physical fitness can be objectively established. The main aim of the research reported in this thesis is to examine the feasibility, the validity and the reliability of physical fitness tests in individuals with SPIMD.

This research addresses the following research questions:

1 Are body composition measurements in participants with SIMD and GMFCS levels I and II feasible and reliable? If so, what are the outcomes of the body composition measurements in these participants [chapter 2]?
2 Are waist circumference measurements in participants with PIMD and GMFCS levels IV and V valid and reliable [chapter 3]?
3 Are the 6 Minute Walking Distance (6MWD) and the adapted Shuttle Run Test (aSRT) in persons with SIMD and GMFCS levels I and II feasible and reliable [chapter 4]?
4 Are the feasibility, validity and test-retest reliability of the adapted Shuttle Run Test (aSRT) protocol performed on a treadmill for persons with SIMD and GMFCS level I sufficient [chapter 5]?
5 Is the modified Berg Balance Scale (mBBS) in persons with SIMD and GMFCS levels I and II feasible and reliable [chapter 6]?
6 Are the Modified Ashworth Scale (MAS) and the Modified Tardieu Scale (MTS) in persons with PIMD and GMFCS levels IV and V feasible and reliable [chapter 7]?
7 What is the level of physical activity of persons with PIMD based on heart rate patterns when compared to ACSM guidelines of healthy physical activity? Differ heart rate patterns according to group differences, days, time of day and is it possible to establish adherent classification in heart rate height and patterns? Is there a relation between heart rate patterns and observed level of activity in persons with PIMD? What is the influence of covariates such as gender, age, and common co-morbidity (motor disabilities, spasticity and sensory disabilities) on heart rate patterns [chapter 8]?

Outline of the thesis

Chapter 2 addresses the feasibility and the test-retest reliability of body composition measures in participants with SIMD. Anthropometric measurements are widely used to reliably quantify body composition and to estimate risks of overweight in both healthy subjects as in patients. However, information about the reliability of anthropometric measurements in participants with severe intellectual and visual disabilities is lacking.

Chapter 3 deals with the validity and reliability of measuring waist circumference in persons with PIMD. Waist circumference as an indicator of abdominal fat is an important predictor of health risks. It is unknown whether waist circumference can be measured validly and reliably when a participant is in a supine position. This assumption however is a critical one when international standards for healthy subjects are applied to persons with PIMD.

Chapter 4 seeks to address the cardiorespiratory component of physical fitness. Cardiorespiratory fitness can be divided into functional exercise and aerobic capacity [26].

Therefore, a study is put forward with the purpose of examining the feasibility and test-retest reliability of both the six-minute walking distance test (6MWD) as an adapted shuttle run test (aSRT) in persons with SIMD.

Chapter 5 examines the feasibility, validity and reliability of the adapted Shuttle Run Test performed on a treadmill, in persons with SIMD.

Sufficient balance is necessary to perform daily activities. Chapter 6 discusses a study with the purpose of determining the feasibility and reliability of the modified Berg Balance Scale (mBBS) in persons with SIMD.

The Modified Ashworth Scale and the Modified Tardieu Scale Muscle examine muscle tonus or level of spasticity. The purpose of the study described in Chapter 7 was to determine the feasibility, the test-retest reliability and interrater reliability of the Modified Ashworth Scale and the Modified Tardieu Scale in persons with PIMD.

Reliably quantifying physical activity levels in persons with SPIMD is important, but also difficult in persons who are not able to walk. Heart rate monitoring may be an indicator of activity levels. Chapter 8 describes heart rate monitoring and heart rate patterns of persons with PIMD. Furthermore, this chapter examines the relative activity of persons with PIMD when compared to ACSM guidelines of healthy physical activity, as well as the correlation between heart rate patterns and level of activity for this specific target group. Finally, the influence of covariates such as gender, age, and common co-morbidity on heart rate height are examined and participants are classified according to heart rate height during physical activity.

Chapter 9 summarizes the main findings and puts them in perspective. Implications and recommendations for further research, methodological analyses and clinical practice are given.

References

1 Schalock R, Brown I, Brown R, Cummins RA, Felce D, Matikka L, Keith KD, Parmenter T. Conceptualization, Measurement, and Application of Quality of Life for Persons With Intellectual Disabilities: Report of an International Panel of Experts. Ment Retard. 2002;40(6):457-470.

2 Walsh PN, Kerr M, Van Schrojenstein Lantman-de Valk HM. Health indicators for people with intellectual disabilities: a European perspective. Eur J Public Health. 2003;13(3 Suppl):47-50

3 World Health Organization (WHO). International Classification of Functioning, Disability and Health. Geneva, 2001

4 Van Splunder J, Stilma JS, Bernsen RM, Evenhuis HM. Prevalence of visual impairment in adults with intellectual disabilities in the Netherlands: cross-sectional study. Eye (Lond) 2006;20(9):1004-10.

5 Van den Broek EG, Janssen CG, van RT, Deen L. Visual impairments in people with severe and profound multiple disabilities: an inventory of visual functioning. J Intellect Disabil Res 2006;50(Pt 6):470-5.

6 Kiestra T, De unieke handicap, referentiemodel voor meervoudige beperkingen. 2005, ISBN 9080638420, Scholma Druk, Bedum

7 Evenhuis HM, Sjoukes L, Koot HM, Kooijman AC. Does visual impairment lead to additional disability in adults with intellectual disabilities? J Intellect Disabil Res. 2009;53(1):19-28.

8 Van Schrojenstein Lantman-de Valk HMJ, Metsemakers JFM, Haveman MJ, Crebolder HFJM. Health problems in people with intellectual disability in general practice: A comparative study. Family Practice. 2000;17:405–407

9 Nakken H & Vlaskamp C. A need for a taxonomy for profound intellectual and multiple disabilities. J Policy Practice Int Dis.2007;4;83-7.

10 Carmeli E, Bar-Yossef T, Ariav C, Paz R, Sabbaq H, Levy R. Sensorimotor impairments and strategies in adults with intellectual disabilities. Motor Control. 2008;12(4):348-61.

11 Shinkfield AJ, Sparrow WA, Day RH. Visual discrimination and motor reproduction of movement by individuals with mental retardation. Am J Ment Retard. 1997;102(2):172-81.

12 Bhaumik S, Watson JM, Thorp CF, Tyrer F, Mc Grother CW. Body mass index in adults with intellectual disability: distribution, association and service implications: a population-based prevalence study. J Intellect Disabil Res. 2008;52(Pt 4):287-98.

13 McGuire BE, Daly P, Smyth F. Lifestyle and health behaviours of adults with an intellectual disability. J Intellect Disabil Res. 2007;51(Pt 7):497-510.

14 Frey GC & Chow B. Relationship between BMI, physical fitness, and motor skills in youth with mild intellectual disabilities. Int J Obes. 2006;30:861-867.

15 Knijff-Raeven van AGM, Jansen-Jacobs CCM, Freen PJW, Hoekman J, Maaskant MA. Body Mass Index (BMI) bij mensen met een verstandelijke beperking. Nederlands Tijdschrift voor de Zorg aan mensen met verstandelijke beperkingen. 2005;1:3-17.

16 Lahtinen U, Rintala P, Malin A. Physical performance of individuals with intellectual disability: a 30 year follow up. Adapt Phys Act Q. 2007;24: 125-143.

17 Stevenson RD, Conaway M, Chumlea WC, Rosenbaum, O'Donnell, Samson-Fang L, Stallings VA, Fung EB, Henderson RC, Worley G, Liptak G. Growth and Health in Children With Moderate-to-Severe Cerebral Palsy. Pediatrics. 2006;118:1010-1018.

18 Draheim CC, Williams DP, MCCubbin JA. Prevalence of physical inactivity and recommended physical activity in community-based adults with mental retardation. Ment Retard.

2002;40:436-444.
19 Temple VA, Frey GC, Stanish HI. Physical activity of adults with mental retardation: review and research needs. Am J Health Promotion. 2006;21:2-12.
20 Haveman M, Heller T, Lee L, Maaskant M, Shooshtari S, Strydom A. Major health risks in ageing persons with intellectual disabilities: an overview of recent studies. J Policy Practice Int Dis. 2010;7:59-69.
21 Houwen S, Visscher C, Lemmink KAPM, Hartman E. Motor skill performance of school-age children with visual impairments. Dev Med and Child Neur. 2008;50(2):139-45.
22 Hopkins WG, Gaeta H, Thomas AC, Hill PM. Physical fitness of blind and sighted children. Eur J of Appl Phys Occ Phys. 1987;56: 69-73.
23 Häkkinen A, Holopainen E, Kautiainen H, Sillanpää E, Häkkinen K. Neuromuscular function and balance of prepubertal and pubertal blind and sighted boys. Acta Paediatr. 2006;95(10):1277-83.
24 Seemungal BM, Glasauwer S, Gresty MA, Bronstein AM. Vestibular perception and navigation in the congenitally blind. J of Neurophys. 2007;97(6):4341-56.
25 Palisano R, Hanna SE, Rosenbaum PL, Rusell DJ, Walter SD, Wood EP, Raina PS, Galuppi BE. Validation of a model of Gross Motor Function for Children With Cerebral Palsy. Phys Therapy. 2000;80:974-985.
26 Bouchard C, Shepard RJ, Stephens T. Physical activity, Fitness and Health. 1994, Human Kinetics Publishers, ChampaignIL.
27 Kramer AF, Hahn S, Cohen NJ, Banich MT, McAyley E, Harrison CR, Chason J, Vakil E, Bardell L, Boileau RA, Colcombe A. Ageing, fitness and neurocognitive function. Nature. 1999;400:418-419
28 Sitskoorn, MM. Het plastische brein; invloed van gedrag. De Psycholoog. 2005;40(5): 262-267.
29 Colcombe S, Erickson KI, Raz N, Webb AG, MC Auley EB, Kramer AF. Aerobic fitness reduces brain tissue loss in aging humans. J Gerontol. 2003;58:176-180.
30 Hale L, Bray A, Littmann A. Assessing the balance capacities of people with profound intellectual disabilities who have experienced a fall. J Intellect Disabil Res. 2007;51(Pt 4):260-8.
31 Rosenbaum P, Paneth N, Leviton A. A report: the definition and classification of cerebral palsy. Dev Med Child Neur. 2007;109:8-14.
32 Warms C. Physical Activity Measurement in Persons with chronic and disabling conditions, methods, strategies and issues. Fam Community Health. 2006;29:788-799.
33 Hunter J. Energy costs of wheelchair propulsion by elderly and disabled people. Int J Rehab Res. 1987;50-54.
34 Littlewood RA, Davies PSW, Cleghorn GJ. Grote RH. Physical activity cost in children tbllowing an acquired brain injury—a comparative study. Clin Nutr. 2004;23:99-104.
35 Sirard JR, Pate RR. Physical activity assessment in children and adolescents. Sports Med. 2001;31(6):439-54.
36 Schiariti V, Fayed N, Cieza A, Klassen A, O'Donnell M. Content comparison of health-related quality of life measures for cerebral palsy based on the International Classification of Functioning, Disability and Rehabilitation. 2010;early online 1-10.
37 Capsersen CJ, Powell KE, Christenson GM. Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research. Public Health Reports.

1985;100(2):127-131.
38 Wilder RP, Green JA, Winters KL, Long WB 3rd, Gubler K, Edlich RF. Physical fitness assessment: an update. Division of Spine and Sports Care, The Runner's Clinic, Department of Physical Medicine and Rehabilitation, University of Virginia Health Sciences System, Charlottesville, Virginia, USA.J Long Term Eff Med Implants. 2006;16(2):193-204.
39 American College of Sports Medicine Position Stand. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. Med Sci Sports Exerc. 1998;30:975-91
40 World Health Organization. Ottawa Charter for Health Promotion, 1986.
41 Hilgenkamp TIM, Van Wijck R, Evenhuis HM. Physical fitness in older people with ID—Concept and measuring instruments: A review. Res Dev Dis. 2010;31:1027–1038.
42 Bouchard C & Sheppard RJ. Physical activity, fitness, and health: The model and key concepts physical activity, fitness and health, international proceedings and consensus statement. 1994, Champain: Human Kinetics Publishers.

Chapter 2

Feasibility and reliability of body composition measurements in adults with severe intellectual and sensory disabilities

A. Waninge
W. van der Weide
I. J. Evenhuis
R. van Wijck
C.P. van der Schans

Journal of Intellectual Disability Research 2009; 53(4):377-388.

Abstract

Background Anthropometric measurements are widely used to reliably quantify body composition and to estimate risks of overweight in healthy subjects and in patients. However, information about the reliability of anthropometric measurements in subjects with severe intellectual and sensory disabilities is lacking.

Objective The purpose of this study was to determine the feasibility and the test–retest reliability of body composition measures in subjects with severe intellectual and sensory disabilities.

Method The study population consisted of 45 subjects with severe intellectual and sensory disabilities. Body mass index, waist circumference, skinfolds and tibia length were measured. Reliability was assessed by Wilcoxon signed rank test, limits of agreement (LOA) and intraclass correlation coefficients. The outcomes were compared with values provided by theWorld Health Organization.

Results There were no significant differences between test and retest ($P < 0.05$). For the skinfold measurements, however, the LOA was insufficient. Intraclass correlation coefficients for all variables, except skinfold measurements, were 0.90 or above.

Conclusion Test–retest reliability and feasibility for all measurements are acceptable in subjects with severe intellectual and sensory disabilities. Skinfold measurements, however, could not be reliably performed in these subjects. Measuring tibia length and using the determined formula to calculate body height from tibia length is a reliable alternative for measuring body height. Although measuring the body height of subjects with severe disabilities was feasible, measuring tibia length was more feasible.

Introduction

Physical fitness and health are related according to the Toronto model [1], in the sense that a good physical fitness may reduce health risks [2, 3]. Health can be defined as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity [World Health Organization (WHO) 4, 5]. In addition, health is considered a resource for everyday life, not the objective of living. Health is a positive concept emphasising social and personal resources, as well as physical capacities [6]. The American College of Sports Medicine [(ACSM), 7] gives the following definition of health-related physical fitness: 'Health related physical fitness is defined as a set of attributes that people have or achieve that relates to the ability to perform physical activity'.

In the ACSM guidelines [8], body composition is defined as a component of health-related physical fitness; this implies that assessment of health-related physical fitness includes measures of body composition [8]. Higher body weights are associated with decrease in health [9]: being obese or overweight substantially increases the risk of morbidity of diseases, like heart and vascular diseases, type 2 diabetes, and respiratory problems [10]. In the Netherlands, over 40% of adults with an intellectual disability (ID) have been shown to be overweight [11]. This figure is similar in other countries [12, 13]. Reliable measurements are essential in order to prevent these individuals from becoming overweight or to reduce the weight of those already overweight.

Anthropometry provides techniques for assessing the size, proportions and composition of the human body; these techniques are universally applicable, inexpensive and non-invasive [14]. To assess an individual's body composition, body length, body weight, waist circumference, skinfold measurement and bioelectrical impedance tests are used [15].

If height cannot be measured, it can be estimated with alternative height measurements such as tibia length, ulna length, knee height or demispan, described by the 'MUST' Explanatory Booklet [16]. Hogan [17] described knee height, Madden [18] ulna length and Weinbrenner [19] demi-span as alternative measurements. Long bone length is known to be the best indicator of stature [20]. Moreover, ulna and tibia length are preferred, because measurements of knee height or demispan may be influenced by deformation of the included joints: the ankle joint in measuring knee height and the shoulder, elbow, wrist and finger joints in measuring demispan. Because of ease of measurement and low cost, tibia length has been advocated by Stevenson [21] as the proxy measurement of choice in mobility-impaired subjects. Duyar & Pelin [20] advised when estimating height based on tibia length, the individual's general stature category should be taken into consideration, and group specific formulae should be used for short and tall subjects.

Body mass index (BMI) provides a more accurate measure of total body fat than body weight alone [15]. The correlation between BMI and body fat content is fairly strong; however, this correlation varies according to gender, race and age [22, 23]. BMI has some limitations: BMI may overestimate body fat in very muscular people and underestimate body fat in some underweight people, who have lost lean tissue, such as the elderly [15].

Another means of assessing body fat content is through waist circumference. Waist circumference as an indicator of abdominal fat, is an important predictor of health risks [15] like heart and vascular diseases and type 2 diabetes [24, 25]. According to the study of Nadas [26], the intra-observer and inter-observer differences in repeated measurements of waist circumference are small when expressed in absolute values.

Some publications regard skinfold thickness as a better predictor of high body fat content

in adults than BMI [27]. Thus, in addition to BMI and waist circumference, it is important to use an additional method to assess body composition, such as skinfold measurements. The reliability of waist circumference and skinfold measurements was examined by Bemben [28] in men aged 20–74. For lean, healthy individuals, most techniques appeared to provide accurate values, but as individuals age there is more discrepancy between the methods. If individuals are frail or not mobile, anthropometry can be used as long as its limitations are noted [28]. Stevenson et al [29] described the reliability of weight, tibia length and skinfold measurements. These authors described that reliability was comparable with other published reports [30] in children with CP. Body composition measurements are widely used in healthy subjects and in patients [31, 15, 32, 27, 26]. In subjects with mild ID, prevalence of overweight and obesity is described among others by Bhaumik et al, Emerson, Melville et al and, Merriman et al [33, 34, 35, 36], using BMI. Furthermore, validity of measurements of BMI, waist-to-hip-ratio and skinfolds in people with learning disabilities was examined by Rimmer [37].

To date, however, no available data exist on the feasibility and reliability of performing these measurements in persons with severe or profound intellectual and sensory disabilities (SIMD). The feasibility and reliability of measuring the body composition of these individuals, however, may be less than that in other subjects, because these persons with severe or profound ID may have an intellectual level of a young child [International Association for the Scientific Study of Intellectual Disabilities (IASSID); 38], may not understand much of their environment, and may be blind or partially sighted and thus cannot see their environment. They are completely dependent on their caregivers and not accustomed to the above-mentioned assessments. Other potential confounding factors include motivational problems, agitation, anxiety and misunderstanding. For example, some are unable to stand up against a wall, whereas others do not understand why they feel a pinch during skinfold measurements. Measuring body composition is very relevant, because these subjects may suffer from inactivity and have increased risk for obesity [39, 40].

The purpose of this study was (1) to determine the feasibility of performing body composition measurements on participants with severe intellectual and sensory disabilities; (2) to determine the test–retest reliability of measuring body composition variables in these participants; and (3) to describe the body composition of these participants.

Materials and methods

Subjects

Participants were classified according to an adapted Gross Motor Function Classification System [(GMFCS), 41], a five-level system used to classify the severity of motor abilities in people with mental and physical disabilities. For example, participants having a 'level 1' classification can generally walk without restrictions but tend to be limited in some more advanced motor skills. Participants with a 'level 5' classification have generally very limited mobility, even with the use of assistive technology. These participants always use a wheelchair.

The original GMFCS was adapted for two reasons:

- In the study population, some participants had better motor skills than those outlined for GMFCS level 1. Thus, we added a level 0 to the classification system; and
- Most of the participants had to deal with impaired vision, and as a result they could not jump and run spontaneously. If a participant spontaneously increased his speed during walking, instead of jumping and running, the participant was classified as GMFCS level 1. The adapted version of the

GMFCS was presented to the investigator, who translated the original version of the GMFCS into Dutch [41] and he concluded that the adaptations did not influence the reliability of the system. The participants were recruited from 'De Brink', a residential care facility in the Netherlands, in which 200 persons with severe of profound intellectual and sensory disabilities live. Moreover, in 65% they are suffering from motor disabilities as well. We asked the representatives of 92 participants a written permission for the subjects to participate in this study. Eighty representatives gave permission. After informed consent was obtained, we screened these participants based on the examination findings of a physician specialized in mental disabilities and of a behavior scholar and excluded five participants. Another eight participants were excluded because they did not live at the centre for people with severe intellectual and sensory disabilities where the tests were performed. Twenty-two participants were excluded because they presented with exclusion criteria (see below) at the time the measurements were being performed (Fig. 1).

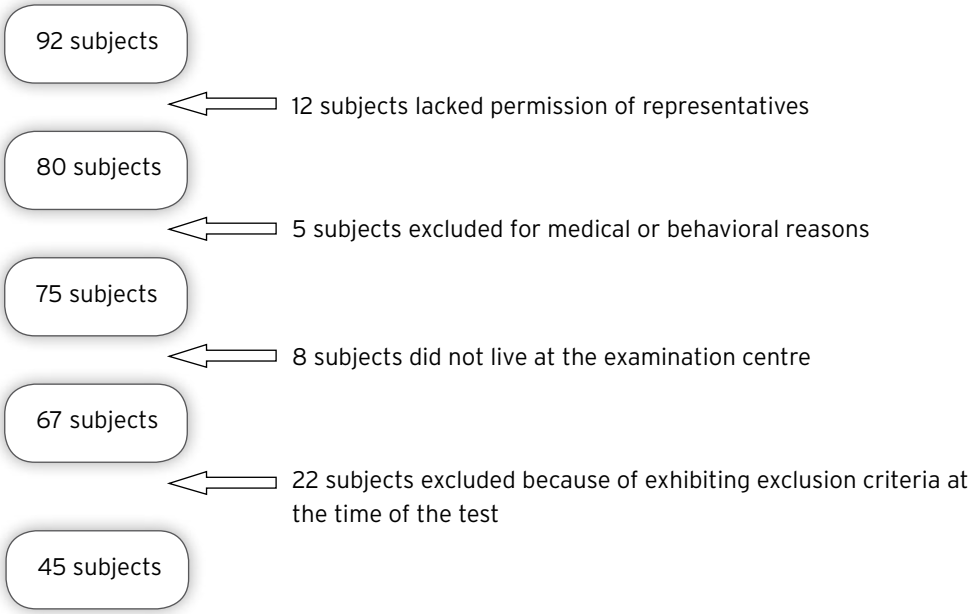


Figure 1. Inclusion steps

In all, 45 participants participated in this study: 17 were female and 28 were male. The mean (SD) age of the men was 38 (11) years and of the women was 44 (10) years. Five participants were classified as GMFCS level 0, 21 participants as GMFCS level 1 and 19 participants as GMFCS level 2. Eighty-nine percent (40) of the participants had severe ID and 11% (5) had profound ID, according to the classification scheme of the IASSID. Most of the participants also had impaired vision. According to WHO guidelines [42], 55% (25) of the clients were severely partially sighted, 38% (17) were partially sighted and 7% (3) were slightly limited in sight. Most participants had impaired motor abilities: 64% (29) had orthopaedic defects. In addition, 29% (13) of the participants had slight hearing problems, 9% (4) had loss of hearing and 4% (2) had severe loss of hearing or were completely deaf.

Study design

Forty-five participants were measured twice. There was 1 week between the test and the retest and both measurements were conducted at the same time of the day. Food before the test–retest, defecation before the test–retest and the attendant of the test–retest were noted so that we could check if these factors influenced whether the tests could not be reliably performed.

Ethical statement

The study was performed in agreement with the guidelines of the Helsinki Declaration as revised in 1975. Permission to carry out the study was obtained from a institutional ethics committee. Informed consent was obtained from representatives of the participants, because all participants were unable to give consent. The measurements were performed in accordance with the behavioral code section entitled 'Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans' [43]. Consistent distress or unhappiness was interpreted as a sign of lack of assent and further participation in the study was reconsidered.

Measures and protocols

All measurements took place around the swimming pool at the institution. This location was chosen because this was thought to be a relaxing environment for the participants. Three testers, a dietary therapist, a physical therapist and a physical therapy student took the measurements after an appropriate training (three times) with two together. Participants were excluded from the study if they exhibited any of the following exclusion criteria at the time of the measurements: psychoses, depression, or other severe psychological problems, or somatic diseases, which were defined as chronic diseases and/or diseases that do not resolve in the short term (e.g. osteoarthritis, osteoporosis, pneumonia, etc.). Participants were also excluded for the following reasons: general illness or fever; taking antibiotics; worsening of asthma, epilepsy (recent insult or epileptic fits), fresh wound(s)/bruise(s) or other factors causing pain during movement; or stress due to the participants' behaviour just before the measurement date.

Body height

The participant was asked to remove his shoes, to stand with his back against a wall and to place his feet flat on the ground such that the back of his heels made contact with the wall. The participant had to stand straight and look forward. When the participant was standing correctly, we determined body length by sliding the measuring tape upward from the ground towards the participant's head (Seca height meter 202, accurate at the 0.1 cm level, Hamburg, Germany). The length was noted in centimeters (cm). When a participant was unable to stand against the wall properly, he was asked to lie in a supine position or to lie on one side with stretched legs for body length to be measured. To measure a participant in a supine position, we drew two horizontal lines – one touching the top of his head and one touching the bottom of his heels – and measured the distance between the two lines. To measure a participant lying on his side, we measured the following distances: (1) top of head-to-cervical spine; (2) cervical spine-to-os sacrum; and (3) os sacrum-to-bottom of heel. We summed these three distances to obtain the body length measurement. For both positions, supine and lying on one's side, it was important to follow the body lines instead of measuring the shortest distance. No protocols from previous studies measuring height in supine position are known. However, our measurer skills are based on 20 years dietary and physiotherapeutic experience in handling participants from this target group.

Body weight

To determine the body weight, we instructed the participant to remove his shoes, to wear only swimming clothes and to stand on an electronic calibrated gauged pair of scales (Weigh plateau for wheelchairs PM-9050, Lopital Nederland BV, Oisterwijk, the Netherlands). When a participant was unable to stand independently without moving (e.g. because of anxiety), the participant was weighed in a wheelchair. The weight (kg) of the participant was then calculated according to the following formula: body weight = measured weight – mass of wheelchair.

BMI

The following formula was used to calculate BMI:
BMI = (body weight, kg)/(body length, m)2.

Waist circumference

We used a measuring tape (Seca 201 tape measurer, accurate at the 0.1 cm level, Hamburg, Germany) to determine waist circumference. Waist circumference was measured at the point located halfway between the crista iliaca and the tenth rib. We took two measurements, one as the participant breathed in and one as he breathed out. The average of these two values was used for analysis. We deviate from The International Society for the Advancement of Kinanthropometry procedure, because in our participants a normal expiration in these circumstances may be difficult to recognize.

Skinfold measurements

Skinfolds were measured in mm at four sites according to the guidelines of the ACSM [8] and Harpenden Skinfold Calliper (Model: HSK-BI, Baty International, West Sussex, UK) was used. The participant was asked to stand straight for all measurements, and skinfolds on the right side were measured twice. The average of the two measurements was used for analysis. Triceps and biceps skinfolds were measured at the midpoint between the acromion border and the proximal border of the olecranon. Subscapular skinfolds were measured by palpating the participants' angulus inferior scapulae and pinching the skinfold located just lateral to and under the angulus inferior at an angle of 45 degrees (with the spine). Suprailiac skinfolds were measured by palpating the crista iliaca of the pelvis and pinching the skinfold just before the top of the crista iliaca.

Tibia length

The participant was asked to sit in a chair with his knees flexed at 90 degrees. Next, we palpated the medial malleolus and the proximal end of the tibia, and then measured the distance (cm) between the distal border of the medial malleolus and the proximal end of the tibia with a measuring tape.

Data analysis

The data were analysed using spss 14.0.

Feasibility

To determine feasibility, we compared the number of successful measurements with the total number of measurements. The feasibility was considered to be sufficient when 95% of the measurements were successful.

Reliability

First, to determine whether significant differences between measurements 1 and 2 exist, we analysed the differences using the Wilcoxon signed rank test. Wilcoxon signed rank tests were used because the data were not normally distributed. The level of statistical significance was set at 5%. Limits of agreement (LOA) between two measurements of the same variables were calculated according to the procedure described by Bland & Altman [44]. The LOA is considered to be an indicator of reliability. LOAs are expressed in units and as a percentage of the mean of the first measurement. Measurements were considered reliable when the LOA was less than 10% of the mean of the first measurement.

Finally, the intraclass correlation coefficients (ICC two-way random, absolute agreement) of measurements 1 and 2 of the same variables were computed. Reliability is considered acceptable when the ICC value is greater than 0.80 and the 95% confidence interval is 0.04 or less. To compare the reliability determined in the present study with those from other studies, we made similar calculations of similar variables: the standard error of measurement (SEM = SD/√n); the coefficient of variation 1 (CV1 = SD/mean x 100%); the technical error (TE = $\sqrt{Sd^2/2n}$, where d is the difference between paired measures of n subjects); and the coefficient of variation 2 (CV2 = 100 x TE/mean of measures taken).

Calculation of height

The long bone length/height ratio has been shown to vary among populations [45] and it is known that this ratio does vary to some degree with differences in stature [46, 47]. To address these differences, specific formulae have been generated for certain populations [48, 49, 50, 51, 52, 53]. For that reason, we used the formula to estimate a subject's height from his tibia length of Stevenson [54]. The calculated height data were compared with the actual measured height using Wilcoxon rank tests and ICC (two-way random, total agreement). Also, LOAs between two measurements of the same variables were calculated according to the procedure described by Bland & Altman [44]. If necessary, a linear regression analysis was performed to determine the most appropriate equation for calculating a subject's height from tibia length.

Body weight status

If feasibility and reliability were acceptable, the outcomes were compared with the normal WHO values [55, 56, 32].

Results

Reliability

Table 1 summarizes the means and standard deviations of body length, body weight, BMI, waist circumference, and skinfold measurements, the results of Wilcoxon signed rank test, LOA, ICC, and LOA expressed as a percentage of the means. There were no significant differences between measurements 1 and 2. The LOAs expressed as a percentage of the means for the skinfold measurements were more than 10%, whereas the LOAs for all the other variables were less than 10%. Intraclass correlation coefficients for all variables, except those for biceps, subscapular and suprailiac skinfold measurements, were 0.90 or above.

Table 1 Results of Wilcoxon rank test, limits of agreement (LOA), percentage LOA of mean, and intraclass correlation coefficients (ICC)

	Mean 1(SD)	**Mean 2 (SD)**	**p value Wilcoxon**	**LOA**	**LOA of mean (%)**	**ICC* 95% CI**
Weight (in kg)	63 (12)	63 (12)	0.814	2 * 1.06939	3.2%	0.99 0.99-0.99
Height (in cm)	164 (14)	164 (14)	0.129	2 * 1.75123	2.1%	0.99 0.99-0.99
BMI (in kg/m^2)	23 (3)	23 (3)	0.554	2 * 0.68659	5.8%	0.98 0.98-0.99
Tibia length (in cm)	36(4)	36(4)	0.527	2 * 0.94545	5.2 %	0.99 0.98-0.99
Waist circumference (in mm)	85 (9)	85 (9)	0.372	2 * 2.27459	5.3%	0.97 0.95-0.98
Skinfold Biceps (in mm)	10 (5)	10 (6)	0.468	2 * 4.02861	80%	0.86 0.75-0.93
Skinfold Triceps (in mm)	15 (6)	16 (7)	0.957	2 * 3.56180	44%	0.91 0.85-0.96
Skinfold Sub scapular (in mm)	19 (8)	20 (7)	0.258	2 * 4.64008	46%	0.83 0.80-0.95
Skinfold Suprailiac (in mm)	20 (7)	20 (7)	0.957	2 * 5.52351	55%	0.89 0.68-0.91
Sum of skinfold measurements (in mm)	66 (22)	65 (21)	0.931	2 * 8.94375	27%	0.91 0.84-0.95

* Two way random, total agreement.

Table 2 shows SEM and CV1 values for waist circumference and for biceps, triceps, subscapular and suprailiac skinfolds. Table 3 lists TE and CV2 values for weight, tibia length, and triceps and subscapular skinfolds.

Table 2 Standard error of measurement (SEM) and coefficient of variation 1 (CV1) values for waist circumference measurements and skinfold measurements*

	SEM Present study	**CV 1 Present study**
Waist measurement	0.340	0.400
Skinfold biceps	0.622	6.212
Skinfold triceps	0.556	3.590
Skinfold subscapular	0.743	3.810
Skinfold suprailiac	0.863	4.300

* SEM = SD/√n; CV1 = SD/mean x 100%

Table 3 Technical error (TE) and coefficient of variation (CV) values for weight, tibia length, and triceps and subscapular skinfolds*

	TE Present study	**CV 2 Present study**
Weight	0.0005	0.0008
Tibia length	0	0
Skinfold triceps	0.03	0.19
Skinfold subscapular	0.07	0.36

*TE = $\sqrt{\sum d^2/2n}$, where d is the difference between paired measures for n subjects; CV = 100 x TE/mean of measures taken

Calculation of height

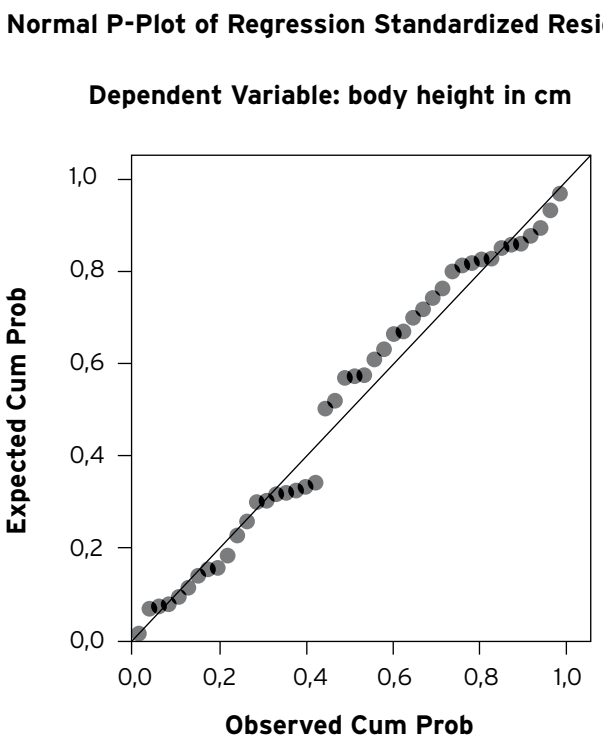
The mean (SD) height calculated with the Stevenson formula [151 (35) cm] was significantly different ($p < 0.01$) from that of the mean (SD) of actual measured heights [164 (11) cm]. Because of this significant difference, we performed a linear regression analysis to identify a more accurate formula for calculating height from tibia length and arrived at the following formula ($p < 0.01$; R/R2: 0.926/0.857; Durbin/Watson: 1.945):

For men, 74.008 + (1.841 x tibia length) + (0.389 x weight) - (3.787 x 0);

For women, 74.008 + (1.841 x tibia length) + (0.389 x weight) - (3.787 x 1).

We used the Wilcoxon rank tests, ICC (2-way random, total agreement) and LOAs to compare height data calculated with the new formula to actual measured height data and found that height calculated with this formula was not significantly different from the actual measured height ($p = 0.953$). Moreover, the LOA was 10 cm, which is 6% of the mean, and the ICC was 0.96. Figure 2 shows that the standardized normal P-P plot of the regression analysis is acceptable.

Figure 2 Normal P-P Plot.



Feasibility

Except for the measurements of skinfolds, which were 82% successful, at least 95% of the remaining measurements were successful.

Body weight status

Tables 4 and 5 show the results of the comparison of BMI and waist measurements according to the WHO [32, 55, 56].

Table 4. BMI interpretation according to WHO values: obese (BMI>30); overweight (25<BMI>30); healthy weight (18.5<BMI>25); underweight (BMI<18.5).

	Whole group	Men	Women
N	45	28	17
Underweight	4%	7%	0
Healthy weight	65%	79%	41%
Overweight	27%	14%	47%
Obese	4%	-	12%
Totaal	100%	100%	100%

BMI, Body Mass Index, in kg/m^2.
WHO, World Health Organization.

Table 5. Waist circumference interpretation according to WHO values: abdominal obese (waist > 102 for men, waist > 88 for women); risk weight (96 > waist < 102 for men, 80 > waist < 88 for women); healthy weight (88 > waist < 96 for men, 76 > waist < 80 for women); underweight (waist < 88 for men, waist < 76 for women).

	Whole group	Men	Women
N	45	28	17
Underweight	16%	25%	-
Healthy weight	42%	61%	12%
Risk weight	24%	11%	47%
Abdominal obese	18%	3%	41%
Totaal	100%	100%	100%

Waist circumference in cm.
WHO, World Health Organization.

Discussion

The results of our study show that measurements such as body height, body weight, waist circumference and tibia length can reliably be performed in participants with severe intellectual and sensory disabilities (SIMD).

Feasibility and reliability of the measurements depended partly on the motivation of the attendant and participant. The environment and the attitude of the attendant can influence a participant's state of mind. However, when a participant is stressed and moves a lot, it is difficult to take a correct measurement. When a participant is relaxed, the attendant has more time to read the measurement value, and thus the measurement will be more accurate. To measure body height, the examiner has to determine whether a participant is standing correctly, because the participant is unaware of his stance. The measurement process must follow the protocol, so the attendant must check that the participant's feet are flat on the ground, that the back of his heels contact the wall and that he is standing straight and is looking forward. This process can be very difficult for the attendant, because it is often hard for a participant to stand still for a few seconds in the correct position. For this reason, we sought another way of determining participant's height by calculating body height from tibia length. To accurately measure tibia length, an attendant must have sufficient knowledge of human anatomy. We found that the feasibility of obtaining accurate measurements from tibia lengths is much better, because the participant is allowed to sit on a chair.

We experienced the most problems in performing skinfold measurements. During the measurement, the participant feels a pinch but does not understand why he or she is being pinched. Hence, at that moment, the participant becomes agitated and starts moving. This restricts measurement, because as soon as a participant feels the pinch, it takes 2 s before it is possible to read the correct value. When the subject is unable to stand still, it is almost impossible to take an accurate measurement. The skinfold measurement process also caused an unacceptable amount of stress to most of the participants. Furthermore, the LOAs expressed as a percentage of the mean skinfold values show that the skinfold measurement accuracy was unacceptable.

The reliability of body weight, body height, waist circumference, skinfolds and tibia length measurements of the present study is comparable to the reliability of similar measurements reported in other studies. This is considered to be a good result because of the complexity of obtaining measurements in this study population. In the study of Bemben et al. (1998), the reliability of waist circumference measurements and skinfold measurements was examined by determining the standard errors of measurement and coefficients of variation. Our waist circumference measurements (SEM/CV: Bemben et al. [28], 0.590/0.72; the present study, 0.340/0.400) and suprailiac skinfold measurements (SEM/CV: Bemben [28], 3.120/20.73; present study, 0.863/4.3) were more accurate than those reported by Bemben [28]. However, Bemben's [28] biceps, triceps and subscapular skinfold measurements are more accurate than our measurements (SEM: Bemben [28], 0.470, 0.420, 0.590, respectively; present study, 0.622, 0.556, 0.743, respectively).

In the study of Stevenson et al. (2006), the reliability of anthropometric measurements was examined by determining the technical error and the coefficients of variation. By comparing their calculations, we found that our weight, tibia length and skinfold (triceps and subscapular) measurements are more accurate (TE: Stevenson [29], 0.08, 0.22, 0.6, 0.51, respectively;

present study, 0.0005, 0, 0.03, 0.07, respectively). In the study of Prince [57], the ICC of waist circumference was 0.99 ($p < 0.0001$) and LOAs from -5.5 to 6.7 cm was 6.1 cm. In our study, the intraclass correlation was similar. However, LOA was 4.4 cm, indicating that our measures were somewhat more sensitive for monitoring individual changes. The study of Nadas [26] examined intra-observer and inter-observer variability in waist circumference measurements and BMI. In their study, the difference of the means of BMI measurement 1 and 2 was 0.02 kg/m^2, and the absolute average difference of the BMI was 0.292 kg/m^2. In our study, the difference of these two means of the BMI was 0.10 kg/m^2, and the absolute difference between BMI values was 0.687 kg/m^2, which is less reliable, but still acceptable, according to the LOAs expressed as a percentage of the means.

The results of the present study also demonstrated that a considerable number of participants with SIMD are overweight or obese, and are therefore at risk for developing health problems. According to the BMI and waist measurements, more of the men than women had a healthy weight. Thus, the women in the study population were at a higher risk for developing health problems compared with the men. Based on BMI values, 10% of the female subjects were obese and 39% were abdominal obese, while 0% of the male clients were obese and only 7% were abdominal obese.

Conclusions

Test–retest reliability and feasibility for all measurements are acceptable in participants with SIMD. However, skinfold measurements could not be reliably performed in these subjects. Measuring tibia length and using the determined formula to calculate body height from tibia length is a reliable alternative for measuring body height. Although the feasibility of performing body height measurements as outlined in our protocol was acceptable, the feasibility of performing tibia length measurements was much better. Assessing body fat composition in adults with SIMD through skinfold measurements is not recommended. Furthermore, our results indicate that this study population has a considerable number of participants that are overweight or obese.

Acknowledgements

The research was financed by Hanze University Groningen, the Brink, and with funding from the Regional Action-and-Attention Knowledge Circulation. The authors kindly acknowledge and thank the participants for their participation in this study, their representatives for given permission to this, Ms J. Kramer and the caregivers of 'De Brink' for assistance with the measurements.

References

1 Bouchard C, Shepard RJ, Stephens T, Physical activity, Fitness and Health. 1994, Human Kinetics Publishers, Champaign, IL.

2 Kramer AF, Hahn S, Cohen NJ, Banich MT, McAyley E, Harrison CR, Chason J, Vakil E, Bardell L, Boileau RA, Colcombe A. Ageing, fitness and neurocognitive function. Nature. 1999;400:418-419.

3 Sitskoorn MM. Het plastische brein; invloed van gedrag. De Psycholoog. 2005;40(5):262-267.

4 World Health Organization. Health for all, 1979.

5 World Health Organization. WHO Chr. 1980;34(2):80.

6 World Health Organization. Ottawa Charter for Health Promotion, 1986.

7 Wilder RP, Green JA, Winters KL, Long WB 3rd, Gubler K, Edlich RF. Physical fitness assessment: an update. Division of Spine and Sports Care, The Runner's Clinic, Department of Physical Medicine and Rehabilitation, University of Virginia Health Sciences System, Charlottesville, Virginia, USA. J Long Term Eff Med Implants. 2006;16(2):193-204.

8 Amstrong LE, Whaley MH, Brubaker PH, Otto RM. ACSM's Guidelines for Exercise Testing and Prescription, 7th ed. Philadelphia, 2005, Pa: Lippincott Williams & Wilkins, American College of sports medicine, guidelines skinfold measurement, Philadelphia, PA.

9 Bouchard C. Overweight, mortality and obesity. In: (Ed. C. Bouchard), 2000: 3-7. Human Kinetics Publishers, Champaign, IL.

10 Bray GA. Overweight, Mortality and Obesity. In: Physical activity and obesity. (Ed. C. Bouchard), 2000, pp. 32-37. Human Kinetics Publishers, Champaign IL.

11 Knijff-Raeven van AGM, Jansen-Jacobs CC M, Freen P JW, Hoekman J, Maaskant MA. Body Mass Index (BMI) bij mensen met een verstandelijke beperking. Nederlands Tijdschrift Voor de Zorg aan Mensen Met Verstandelijke Beperkingen. 2005;1:3-17.

12 Frey GC & Chow B. Relationship between BMI, physical fitness, and motor skills in youth with mild intellectual disabilities. Intern J Obes. 2006;30(5):861-7.

13 Lahtinen U, Rintala P & Malin A. Physical performance of individuals with intellectual disability: a 30 year follow up. Adapt Phys Act Q. 2007;24(2):125-43.

14 World Health Organization. Physical status: The use and interpretation of anthropometry. 1995, WHO Technical Report Series World Health Organization, Geneva.

15 Hark L, Bowman M, Bellini L. Oversight of Nutrition in Clinical Care. In: Medical Nutrition & Disease (eds L. Hark & G. Morrison). 2003, National Health and Nutrition Examination Survey IV.

16 Malnutrition Advisory Group. The Malnutrition Universal Screening Tool [MUST] Explanatory Booklet. 2003, BAPEN Redditch, Worcs.

17 Hogan S E. Knee Height as a Predictor of Recumbent Length for Individuals with Mobility-Impaired Cerebral Palsy. Journal of American College of Nutrition. 1999;18:201–5.

18 Madden AM, Tsikoura T, Stott DJ. The estimation of body height from ulnar length in adults from different ethnic groups. J of Hum Nutr Diet. 2008;15;21(4):394.

19 Weinbrenner T, Vioque J, Barber X, Asensio L. Estimation of height and body mass index from demi-span in elderly individuals. Gerontology. 2006;52(5):275-81.

20 Duyar I & Pelin C. Body height estimation based on tibia length in different stature groups. Am J Phys Anthropol. 2003;122(1):23-27.

21 Stevenson RD. Measurement of Growth in Children with Developmental Disabilities. Dev Med Child Neurol. 1996;38:855-860.
22 Gallagher D, Visser M, Sepúlveda D, Pierson RN, Harris T, Heymsfield SB. How useful is BMI for comparison of body fatness across age, sex and ethnic groups? Am J Epidemiol. 1996;143:228–239.
23 Deurenberg-Yap M, Schmidt G, Staveren WA van, Deurenberg P. The paradox of low body mass index and high body fat percentage among Chinese, Malays and Indians in Singapore. Int J Obes. 2000;24(8):1011-1017.
24 Savva SC, Tornaritis M, Savva ME, Kourides Y, Panagi A, Silikiotou N, Georgiou C, Kafatos A. Waist circumference and waist-to-height ratio are better predictors of cardiovascular disease risk factors in children than body mass index. Int J Obes Rel Metabolic Dis. 2000;24(11):1453-8.
25 Dalton M, Cameron AJ, Zimmet PZ, Shaw JE, Jolley D, Dunstan DW, Welborn TA, Waist circumference, waist-hip ratio and body mass index and their correlation with cardiovascular disease risk factors in Australian adults. J Intern Med. 2003;254(6):555-63.
26 Nadas J, Putz Z, Kolev G, Nagy S, Jermendy G. Intraobserver and interobserver variability of measuring waist circumference. Med Sci Monitor. 2008;14(1):CR 15-18.
27 Nooyens AC, Koppes LL, Visscher TL, Twisk JW, Kemper HC, Schuit AJ, van Mechelen W, Seidell JC. Adolescent skinfold thickness is a better predictor of high body fatness in adults than is body mass index: the Amsterdam Growth and Health Longitudinal Study. Am J Clin Nutr. 2007;85(6):1533-9.
28 Bemben MG, Massay BH, Bemben DA, Boileau RA, Misner JE. Age-related variability in body composition methods for assessment of percent fat and fat-free mass in men aged 20-74 years. Age and Ageing. 1998;27:147-53.
29 Stevenson RD, Conaway M, Chumlea WC, Rosenbaum P, O'Donnell M, Samson-Fang L, Stallings VA, Fung EB, Henderson RC, Worley G, Liptak G. Growth and Health in Children With Moderate-to-Severe Cerebral Palsy. Pediatrics. 2006;118:1010-1018.
30 Spender QW, Cronk CE, Charney EB, Stallings VA. Assessment of linear growth of children in cerebral palsy: use of alternative measures to height or length. Dev Med Child Neurol. 1989;31:206–214.
31 Mei Z, Grummer-Strawn LM, Pietrobelli A, Goulding A, Goran MI, Dietz WH. Validity of body mass index compared with other body-composition screening indexes for the assessment of body fatness in children and adolescents. Am J Clinic Nutr. 2002;7597–985.
32 World Health Organization, "Waist circumference", 2006b, World Health Organization.
33 Bhaumik S, Watson JM, Thorp CF, Tyrer F, Mc Grother CW. Body mass index in adults with intellectual disability: distribution, association and service implications: a population-based prevalence study. J Intellect Disabil Res. 2008;52(4):287-98.
34 Emerson E. Underweight, obesity and exercise among adults with intellectual disabilities in supported accommodation in Northern England. J Intellect Disabil Res. 2005;49(2):134-43
35 Melville CA, Cooper SA, McGrother CW, Thorp CF, Collacott R. Obesity in adults with Down syndrome: a case-control study. J Int Dis Res. 2005;49(2):125-33.
36 Merriman S, Haw C, Kirk J, Stubbs J. Risk factors for coronary heart disease among inpatients who have mild intellectual disability and mental illness. J Intellect Disabil Res.. 2005;49(5):309-16.

37 Rimmer J, Kelly LE, Rosentswieg J. Accuracy of Anthropometric Equations Estimating Body Composition of Mentally Retarded Adults. Am J Ment Retard. 1987;91(6):626-632.
38 Luckacsson R, Borthwick-Duffy S, Buntinx W, Coulter D, Craig P, Reeve A, Shalock R, Snell M, Spitalnik D, Spreat S, Tasse, M. Mental Retardation: Definition, Classification and Systems of Supports. 2002, Washington American Association on Mental Retardation.
39 Draheim CC, Williams DP, McCubbin JA Prevalence of physical inactivity and recommended physical activity in community-based adults with mental retardation. Ment Retard. 2002;40(6):436-44.
40 Temple VA, Frey GC, Stanish HI. Physical activity of adults with mental retardation: review and research needs. Am J Health Promot. 2006;21(1):2-12.
41 Palisano R, Hanna SE, Rosenbaum PL, Rusell DJ, Walter SD, Wood EP, Raina PS, Galuppi BE. Validation of a model of Gross Motor Function for Children With Cerebral Palsy. Physical Therapy. 2000;80(10):974-985.
41 Gorter JW. Gross Motor Function Classification System (Dutch translation), 2001, Revalidatiecentrum De Hoogstraat Utrecht.
42 World Health Organization. International classification of functioning, disability and health. 2001, Geneva
43 Dutch Society for Doctors in the Care for people with an Intellectual Disability (NVAZ). Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans, behavioural code for doctors in the assessment of resistance among people with an intellectual disability, 1999.
44 Bland JM & Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986;8:307-10.
45 Meadows Jantz L & Jantz RL. Secular change in long bone length and proportion in the United States, 1800–1970. Am J Phys Anthropol. 1999;110:57–67.
46 Meadows L & Jantz RL. Allometric secular change in the long bones from the 1880s to the present. J For Sci. 1995;40:762–767.
47 Konigsberg LW, Hens SM, Jantz LM, Jungers WL. Stature estimation and calibration: Bayesian and maximum likelihood perspectives in physical anthropology. Yearbook of Physical Anthropology. 1998;41:65–92.
48 Breitinger E. Zur Berechnung der Korperhohe aus den langen Gliedmassenknochen. Anthropologischer Anzeiger. 1937;14:249–274.
49 Trotter M & Gleser G. Estimation of stature from long bones of American whites and Negroes. Am J Phys Anthropol. 1952;10:463–514.
50 Trotter M & Gleser G. A re-evaluation of estimation of stature based on measurements of stature taken during life and long bones after death. Am J Phys Anthropol. 1958;16:79–123.
51 Olivier G. Practical anthropology. 1969, Charles C. Thomas Springfield, IL.
52 Formicola V & Franceschi M. Regression equations for estimating stature from long bones of early Holocene European samples. Am J Phys Anthropol. 1996;100:83–88.
53 De Mendonca MC. Estimation of height from the length of long bones in a Portuguese adult population. Am J Phys Anthropol. 2000;112:39–48.
54 Stevenson RD. Measurement of Growth in Children with Developmental Disabilities. Dev Med Child Neurol. 1996;38:855-860.

55 World Health Organization, BMI classification, 2006, World Health Organization.
56 World Health Organization, Pan American Health Organization. Protocol for the Nutritional Management of Obesity, Diabetes and Hypertension in the Caribbean, 2004, Caribbean Food and Nutrition Institute Jamaica and PAHO/WHO Office of Caribbean Program Coordination, Barbados.
57 Prince SA, Janssen I, Tranmer JE. Self-Measured Waist Circumference in Older Patients With Heart Failure: A Study of Validity and Reliability Using a Myo Tape(R), J Cardiopulm Rehabil Prev. 2008;28(1):43-47.

Chapter 3

Measuring waist circumference in disabled adults.

A. Waninge
K. A. M. Ligthart
J. Kramer
S. Hoeve
C.P. van der Schans
H.H. Haisma

Research in Developmental Disabilities 31(3):839-847.
Reproduced by courtesy of Elsevier

Abstract

To date, it is unknown whether waist circumference can be measured validly and reliably when a subject is in a supine position. This issue is relevant when international standards for healthy participants are applied to persons with severe intellectual, sensory, and motor disabilities. Thus, the aims of our study were (1) to determine the validity of waist circumference measurements obtained in a supine position, (2) to formulate an equation that predicts standing waist circumference from measurements obtained in a supine position, and (3) to determine the reliability of measuring waist circumference in persons with severe intellectual, sensory, and motor disabilities. First, we performed a validity study in 160 healthy participants, in which we compared waist circumference obtained in standing and supine positions. We also conducted a test-retest study in 43 participants with severe intellectual, sensory, and motor disabilities, in which we measured the waist circumference with participants in the supine position. Validity was assessed with paired t-test and Wilcoxon signed rank test. A prediction equation was estimated with multiple regression analysis. Reliability was assessed by Wilcoxon signed rank test, limits of agreement (LOA), and intraclass correlation coefficients (ICC). Paired t-test and Wilcoxon signed rank test revealed significant differences between standing and supine waist circumference measurements. We formulated an equation to predict waist circumference (R^2=0.964, $p<0.001$). There were no significant differences between test and retest waist circumference values in disabled participants (p=0.208; Wilcoxon signed rank test). The LOA was 6.36 cm, indicating a considerable natural variation at the individual level. ICC was .98 ($p<0.001$). We found that the validity of supine waist circumference is biased towards higher values (1.5 cm) of standing waist circumference. However, standing waist circumference can be predicted from supine measurements using a simple prediction equation. This equation allows the comparison of supine measurements of disabled persons with the international standards. Supine waist circumference can be reliably measured in participants with severe intellectual, sensory, and motor disabilities.

Introduction

Children and adults with severe generalized cerebral palsy (CP) and intellectual disability have an increased risk for malnutrition [1]. This is a consequence of an altered energy metabolism [2] in combination with feeding difficulties such as gastro-esophageal reflux and dysphagia [3]. Many of these subjects need to be fed by stomach tube. Malnutrition is associated with poorer health status and limitations in societal participation [4]. On the other hand, 40% of the adults with intellectual disability in the Netherlands [5] and in other countries [6, 7] have been found to be overweight. These adults have increased risk for developing obesity [8, 9] and associated degenerative diseases such as type 2 diabetes.

Anthropometry provides techniques for assessing the size, proportions, and composition of the human body; these techniques are universally applicable, inexpensive, and non-invasive [10]. To assess an individual's body composition, body mass index (BMI, kg/m^2) can be used. The correlation between BMI and body fat content is fairly strong; however, this correlation varies according to gender, race, and age [11, 12]. Furthermore, BMI has some limitations, as it may overestimate body fat in very muscular people and underestimate body fat in some underweight people who have lost lean tissue, such as the elderly [13].

Another means of assessing body fat content is through waist circumference. Waist circumference as an indicator of abdominal fat is an important predictor of health risks [13] such as heart and vascular diseases and type 2 diabetes [14, 15]. BMI and waist circumference are widely used measures in healthy participants and in patients [13, 16, 17, 18]. Pischon et al [19] described that 'both general adiposity and abdominal adiposity are associated with increased morbidity and mortality and support the use of waist circumference in addition to BMI in assessing the risk of death'.

De Brink is a residential care facility in the Netherlands, housing 200 persons with severe or profound intellectual, sensory, and in several cases, motor disabilities (PIMD). In a pilot study, we found that the female residents of De Brink appeared to be at a higher risk for developing health problems compared to male residents [20]. In that study, BMI as well as waist circumference were measured. According to BMI values, 10% of the female participants were obese, while none of the male participants were obese. However, when waist circumference was used as a criterion, 39% of the female and 7% of the male participans were classified as being obese. Other authors also conclude that, if waist circumference is used as the criterion, then the prevalence of obesity among these adults may be significantly greater than as indicated by BMI [21, 22].

Reliable measurements are critical for assessing the nutritional status of patients with intellectual disabilities. Reliable measurements are also required to obtain reliable data on prevalence and to identify participants at risk of becoming overweight or developing malnutrition.

We determined that measuring waist circumference with a tape measure halfway between the tenth rib and the hipbone is feasible and reliable in participants with intellectual and sensory disabilities who are able to stand upright [20]. However, due to severe generalized CP and motor disabilities, e.g., spasticity, many participants with intellectual and sensory disabilities are unable to stand straight or stand at all [23]. In these participants, waist circumference can only be measured with the subject lying in a supine position. This raises the question of whether waist circumference can be measured reliably and validly in a supine position. This issue is particularly relevant when international standards for healthy individuals are applied to disabled persons. Therefore, the purpose of this study, was as follows:

(1) in healthy participants, to determine the validity of waist circumference measurements obtained in participants lying in a supine position (supine waist circumference) by comparing these measurements with waist circumference measurements obtained in the same participants in a standing position (standing waist circumference);
(2) to formulate an equation that predicts standing waist circumference based on supine waist circumference and based on covariates that can influence waist circumference, such as gender, age, BMI, or past pregnancy; and
(3) in participants with severe intellectual, sensory, and motor disabilities, to determine the reliability of measuring waist circumference using a test-retest study design.

Methods

Validity study

Study design

The waist circumference of 160 healthy participants was measured while persons were in a standing position and in a supine position.

Participants

One hundred sixty healthy persons without disabilities served in the validity study, in which we compared waist circumference measurements obtained while the participants were in standing and supine positions. Participants were recruited from a nursing school (students and teachers) and from a research organization where people receive medical examinations. All potential participants received written and spoken information about the study. They were included in the study if informed consent was obtained. The participants had to be able to stand and to lie down. Exclusion criteria were pregnancy and having scars, because these situations might alter the shape of the waist.

To ensure that all ages were represented in the study population, we included both men and women from three age categories: 20-35 years, 35-50 years, and 50-65 years. Similarly, all BMI categories were included in the study.

Ethical statement

The participants of this study gave informed consent.

Measurements

A non-stretchable tape measure (Seca 201 tape measure; Seca, Hamburg, Germany), accurate to the 0.1 cm level, was used to determine waist circumference. Waist circumference was measured at the point located halfway between the crista iliaca and the tenth rib. In healthy participants, measurements were obtained while the participants were in a standing position and in supine position. We took two measurements, one as the participant breathed in and one as he/she breathed out. The average of these two values was used for analysis.

Data analysis

The number of participants required was based on a power analysis using data from a pilot study. In order to detect a statistically significant difference of 1.5 cm between the standing and supine measurements, assuming a standard deviation of 9 cm, the study needed to include at least 160

participants. These calculations assume a type I error (alpha) of 0.05, two-tailed, and a type II error (beta) of 20%; that is, a statistical power of 80%. The data were analyzed using SPSS 14.0.

To determine whether significant differences between supine waist circumference and standing waist circumference exist, we analyzed the differences using both a paired t-test and Wilcoxon signed rank test. Wilcoxon signed rank tests were also used to get a better impression of the distribution of the data. The level of statistical significance was set at 0.05.

Furthermore, limits of agreement (LOA) between supine and standing waist circumference measurements were calculated according to the procedure described by Bland and Altman [24].

Predicting standing waist circumference

To determine whether standing waist circumference can be predicted by using supine waist circumference and to determine the influence of the covariates gender, age, BMI, or past pregnancy, first we performed a simple linear regression analysis of standing waist circumference on each variable separately. Significance ($p<0.05$) and R^2 were estimated for each variable. The normality and the homogeneity of variance of the residuals were checked with a normal P-P plot and a plot of the variance.

Subsequently, a model was built with multiple linear regression by first adding all the significant variables and then removing insignificant variables, starting with those having the highest p-value (backward method). Significance and R^2 were estimated from the model, and the normality and the homogeneity of variance of the residuals were checked with a normal P-P plot and a plot of the variance.

Reliability study

Study design

Forty-three disabled persons were measured at two different times, initially at the test and 1 week later at the retest. For each participant, both measurements were conducted at the same time of the day. We recorded information about food intake before the test and retest measurements, defecation before the test and retest, and the attendant of the test and retest.

Participants

For the reliability study, we asked the representatives of 54 persons with severe or profound intellectual, sensory, and motor disabilities (PIMD) written permission for these persons to participate in our study. Forty-eight representatives gave permission. After informed consent was obtained, we screened the 48 persons based on the examination findings of a physician specialized in mental disabilities and a behavior scholar. The screening exclusion criteria were severe psychological problems or somatic diseases, which were defined as chronic diseases and/or diseases that do not resolve in the short term. Three persons were excluded from the study because they had one of these problems or diseases. The exclusion criteria at the time the measurements were being performed were general illness or fever; taking antibiotics; worsening of asthma, epilepsy (recent insult or epileptic fits); fresh wound(s)/bruise(s) or other factors causing pain during movement; or stress due to the participants behavior just before the measurement date. Two persons were excluded because they exhibited one of these criteria. Figure 1 presents the sampling scheme of persons included in the reliability study.

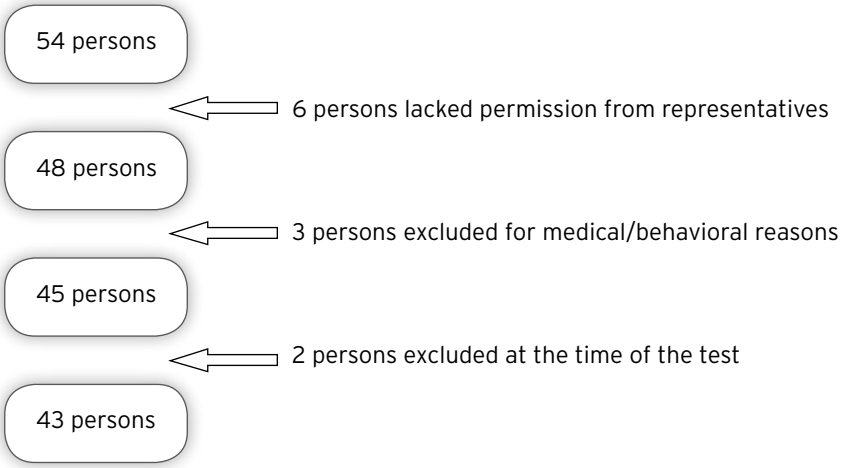


Figure 1. Sampling scheme of subjects included in the reliability study

The participants with PIMD were classified according to the Gross Motor Function Classification System (GMFCS) [25], a five-level system used to classify the severity of motor abilities in people with physical disabilities. For example, persons having a Level I classification can generally walk without restrictions but tend to be limited in some more advanced motor skills. Persons with a Level V classification generally have very limited mobility, even with the use of assistive technology. These persons always use a wheelchair.

Ethical statement

This study was performed in agreement with the guidelines of the Helsinki Declaration, as revised in 1975. Permission to carry out the study was obtained from the institutional ethics committee. Informed consent was obtained from representatives of the participants, because all participants were unable to give consent. The measurements were performed in accordance with the behavioral code section entitled, "Resistance among people with an intellectual disability in the framework of the Governing Medical-Scientific Research Involving Humans Act" [26]. This is a behavioral code for doctors to help them assess the resistance of people with an intellectual disability. The code was drafted by the Dutch Society for Doctors in the Care of People with an Intellectual Disability (NVAZ). Consistent distress or unhappiness was interpreted as a sign of lack of assent, and further participation in the study was reconsidered.

Measurements

A non-stretchable tape measure (Seca 201 tape measure; Seca, Hamburg, Germany), accurate to the 0.1 cm level, was used to determine waist circumference. Waist circumference was measured at the point located halfway between the crista iliaca and the tenth rib, while the disabled participants were in a supine position. We took two measurements, one as the participant breathed in and one as he/she breathed out. The average of these two values was used for analysis. Three testers—a dietary therapist, a physical therapist, and a student—took the measurements after appropriate training (three times).

Data analysis

The data were analyzed using SPSS 14.0. First, to determine whether significant differences between test and retest measurements exist, we analyzed the differences using the Wilcoxon signed rank test. The level of statistical significance was set at 0.05. Limits of agreement (LOA) between two measurements of the same variables were calculated according to the procedure described by Bland and Altman [24]. The LOA is considered to be an indicator of reliability. LOAs are expressed in units and as a percentage of the mean of the first measurement. Measurements were considered reliable when the LOA was less than 10% of the mean of the first measurement. Afterward, the intraclass correlation coefficients (ICC; two-way random, absolute agreement) of test and retest measurements of the same variables were computed. Measurements were considered reliable when the ICC values were greater than 0.80 and the 95% confidence interval (CI) was 0.30 or less. Finally, the test-retest was considered reliable if (1) there were no significant differences between test and retest measurements; (2) LOA was acceptable, as described above; and (3) ICC was acceptable, as described above.

Results

Validity study

The characteristics of the subjects that participated in the validity study are shown in Table 1.

Table 1. Validity study: subject characteristics

Gender			BMI category		
			<25	≥25	Total
Men	Age category	20-35 y	23	7	30
		35-50 y	7	14	21
		50-65 y	13	16	29
		Total	43	37	80
Women	Age category	20-35 y	24	5	29
		35-50 y	13	11	24
		50-65 y	15	12	27
		Total	52	28	80

In all, 160 healthy persons participated in our study, 80 female and 80 male. The mean (SD) age of the men was 42 (15) years and that of the women was 40 (15) years. The mean (SD) BMI of the men was 25 (4) and that of the women was 24 (4).

There were significant differences between standing and supine waist circumference in healthy subjects ($p<0.001$; paired t test). The mean (SD) standing waist circumference was 89.3 (13) cm and the mean (SD) supine waist circumference was 87.8 (12) cm. The Wilcoxon signed rank test ($p<0.001$) showed that in the majority of subjects (n=112) supine waist circumference was lower than standing waist circumference. In 48 subjects (11 men and 37 women),

supine waist circumference was higher than standing waist circumference. In zero subjects, there was no difference. The LOA was 5.34 cm (Figure 2.).

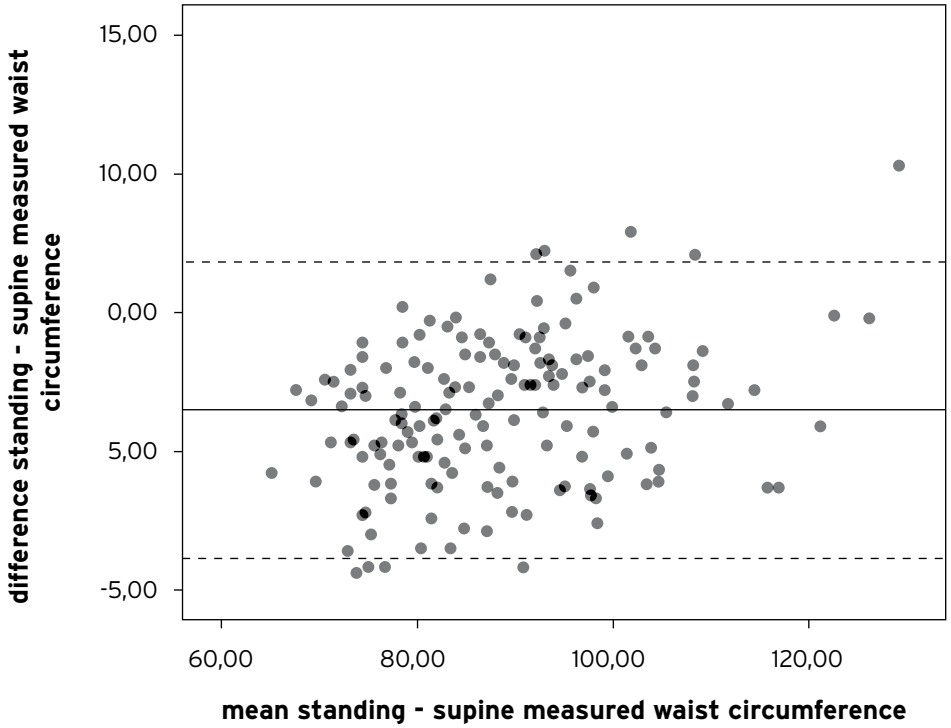


Figure 2. Bland and Altman plot of the differences between standing and supine waist circumference measurements. The mean difference is 1.48±5.34 (LOA) (-3.86; 6.82).

Predicting standing waist circumference

A simple linear regression analysis was performed on standing waist circumference, supine waist circumference, age, BMI, gender, and past pregnancy (Table 2). The normal P-P plots and the plots of the homogeneity of variance residuals showed that there was a normal distribution and homogeneity of variance of the residuals.

Table 2. Simple regression analysis of standing waist circumference using supine waist circumference, gender, age, BMI, and past pregnancy as predictors.*

Simple regression analysis				
Model	Beta	95% CI	p-value	R^2
Waist circumference (supine)	1.044	1.010 to 1.078	<0.001	96%
Gender	-10.184	-13.871 to -6.496	<0.001	16%
Age	0.319	0.192 to 0.446	<0.001	14%
BMI category	17.439	4.399 to 20.478	<0.001	45%
Past pregnancy	-5.205	-9.814 to -0.596	0.027	3%

*R^2, p<0.05, beta, and 95% confidence interval (CI).

A multiple linear regression model showed that BMI (p=0.632), age (p=0.525), and past pregnancy (p=0.084) had no significant influence on waist circumference. Thus, they were removed from the model (Table 3).

Table 3. Multiple regression analysis of standing waist circumference using supine waist circumference, gender, age, BMI, and past pregnancy as predictors.*

Multiple regression analysis				
Model	Beta	95% CI	p-value	R^2
Constant	0.437	-3.049 to 3.996	0.791	
Waist circumference (supine)	1.022	.972 to 1.073	<0.001	
Gender	-1.459	-2.447 to -0.470	0.004	96%
Age	0.10	-0.21 to 0.41	0.525	
BMI category	-0.271	-1.385 to 0.843	0.632	
Past pregnancy	-1.017	-2.171 to 0.137	0.084	

*R^2, p<0.05, beta, and 95% confidence interval (CI).

Standing waist circumference can be predicted by supine waist circumference with a simple correction using the following formula ($p < 0.001$; R/R^2: 0.982/0.964): corrected standing waist circumference = 1.017 – 1.961 x gender + 1.016 x supine waist circumference. The normality and the homogeneity of variance of the residuals were checked with a normal P-P plot and a plot of the variance. The variance of the residuals was found to be homogeneous.

Reliability study

Forty-three individuals in this study. The characteristics of these persons are shown in Table 4.

Table 4. Reliability study: characteristics of the participants

Gender									
		Men				Women			
		Age category							
		20-35	35-50	50-65	Total	20-35	35-50	50-65	Total
Intellectual disability	Severe	4	6	7	17	3	4	4	11
	Profound	4	4	2	10	4	1	0	5
	Total	8	10	9	27	7	5	4	16
Visual impairments	Blind/Severe	5	9	9	23	4	3	3	10
	Partially	3	1	0	4	3	2	1	6
	Total	8	10	9	27	7	5	4	16
Orthopedic defects	Yes	8	9	9	26	7	5	3	15
	No	0	1	0	1	0	0	1	1
	Total	8	10	9	27	7	5	4	16
Cerebral palsy	Yes	6	7	1	14	5	3	0	8
	No	2	3	8	13	2	2	4	8
	Total	8	10	9	27	7	5	4	16
GMFCS level	Level 3	1	1	4	6	2	1	2	5
	Level 4	0	5	4	9	0	2	1	3
	Level 5	7	4	1	12	5	2	1	8
	Total	8	10	9	27	7	5	4	16

GMFCS, Gross Motor Function Classification System.

There were no significant differences between test and retest measurements of waist circumference in disabled subjects (p=0.208; Wilcoxon signed rank test). The mean (SD) of the test measurement was 84 (12) cm, whereas the mean (SD) of the retest measurement was 83 (12) cm. The LOA was 6.36 cm, which is 8% of the mean (Figure 3). ICC was .98; 95% CI, 0.97-0.99 ($p<0.001$).

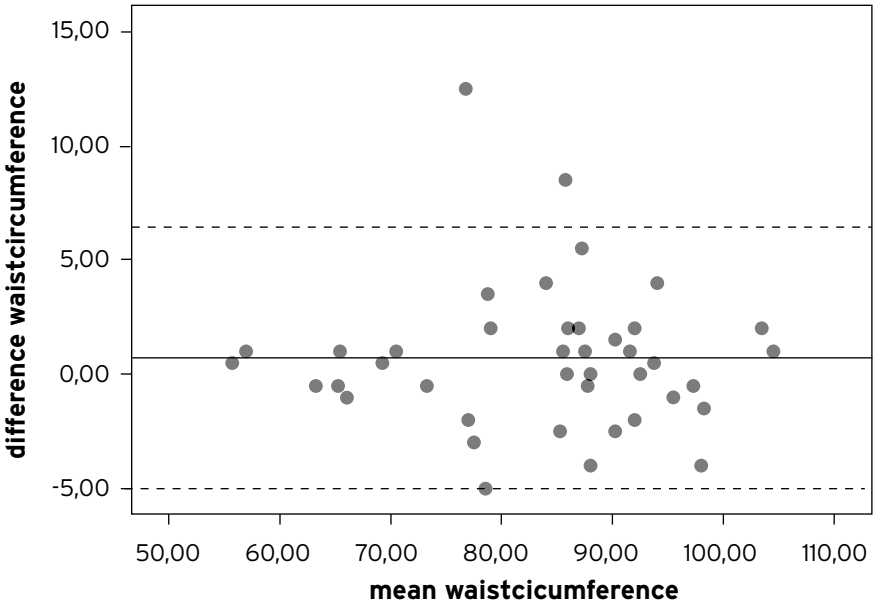


Figure 3. Bland and Altman plot of test and retest supine waist circumference measurements. The mean difference is 0.73 ±6.36 (LOA) (-5.63; 7.09).

Discussion

The purpose of this study was (1) to determine the validity of waist circumference measured in a supine position (supine waist circumference) by comparing these measurements with waist circumference measured in a standing position (standing waist circumference) in healthy participants; (2) to develop an equation to predict standing waist circumference based on supine waist circumference and taking into consideration covariates that can influence waist circumference measurements, such as gender, age, BMI or past pregnancy; and (3) to determine the reliability of measuring waist circumference in participants with severe or profound intellectual, sensory, and motor disabilities using a test-retest study design.

The results of our study show that the validity of supine waist circumference is poor, with higher values (1.5 cm) for standing waist circumference in the majority of healthy participants. This implies that international standards based on unmodified standing measurements from healthy participants cannot be used in disabled persons in whom measurements are conducted in a supine position. We found that age, BMI, and past pregnancy did not influence differences between supine and standing measurements. However, because gender did influence the difference between these two measurements, we formulated a simple equation enabling us to compare the supine measurements obtained from disabled persons with the international standards. Furthermore, we found that measuring waist circumference in a supine position can be reliably performed in participants with PIMD.

Reliable measurements are critical for the assessment of nutritional status in persons with PIMD. These individuals are at risk of becoming either overweight [8, 9] or developing malnutrition [1]. As shown at De Brink, as well as in other studies, women are at higher risk than men for developing health problems caused by overweight [7, 20, 27, 28, 29]. The reliability of the waist circumference measurements obtained in our reliability study is comparable to those

reported in another study [30]. This is considered to be a good result because of the complexity of obtaining measurements in this study population. In the study of Prince et al. [30], the ICC for waist circumference was 0.99 ($p < 0.001$) and LOAs ranged from -5.5 to 6.7 cm, with a mean of 6.1 cm. In our study, the ICC and LOAs were similar: .98 ($p<0.001$) and LOAs ranging from -5.63 to 7.09 (with a mean of 6.36 cm), respectively.

Children and adults who have severe generalized Cerebral Palsy (CP) and intellectual disabilities are often fed by stomach tube [1]. However, tube feeding may improve body weight mainly through fat deposition [31]. Sullivan et al. [32] demonstrated that children with severe CP have relatively low energy expenditure and high body-fat content and highlighted the potential risk of overfeeding with available enteral feeds administered via gastrostomy tube. Therefore, it is necessary to determine the validity of waist circumference measured in a supine position. As far as we know, the validity of these measurements was unknown until now.

A limitation of our study is the fact that the validity study was performed in persons without disabilities. This may influence the outcomes of the equation we formulated for predicting standing waist circumference based on supine waist circumference. However, our results can be applied to a larger group of people, not just persons with PIMD. The equation can also be used to predict standing waist circumference in persons with motor disabilities who are unable to stand.

Another limitation of our study is that the reliability study involved a relatively small number of participants. However, because of the exclusion criteria, there was only a small group of persons with PIMD who were able to participate in the research.

In conclusion, although supine waist circumference can be reliably measured in people with severe or profound intellectual, sensory, and motor disabilities, these measurements cannot be compared with standard waist circumference measures, which are obtained in subjects who are standing. Therefore, a correction equation, such as the one proposed in the present study, is required if such comparisons are to be made.

Acknowledgements

The research was financed by Hanze University Groningen and the Brink.
The authors kindly acknowledge and thank the participants for their participation in this study and their representatives for giving permission to this.
Waist circumference was measured with a non-stretchable tape measurer, Seca Hamburg, Germany.

References

1 Stevenson RD, Conaway M, Chumlea WC, Rosenbaum, O'Donnell, Samson-Fang L, Stallings VA, Fung EB, Henderson RC, Worley G, Liptak G. Growth and Health in Children With Moderate-to-Severe Cerebral Palsy. Pediatrics. 2006;118:1010-1018

2 Veugelers R, Penning C, Van Gulik ME, Tibboel D, Evenhuis HM. Feasibility of bioelectrical impedance analysis in children with a severe generalized cerebral palsy. Nutrition. 2006;22:16–22.

3 Böhmer CJ, Niezen-de Boer MC, Klinkenberg-Knole, EC, Deville WL, Nadorp JH, Meuwissen SG. The prevalence of gastroesophageal reflux disease in institutionalized intellectually disabled individuals. Am J Gastroenterology. 1999;94: 804–10.

4 Samson-Fang L, Fung E, Stallings VA, Conaway M, Worley G, Rosenbaum P, Calvert R, O'Donnell M, Henderson RC, Chumlea WC, Liptak GS, Stevenson RD. Relationship of nutritional status to health and societal participation in children with cerebral palsy. J Pediatrics. 2002;141:637-43.

5 Knijff-Raeven van AGM, Jansen-Jacobs CCM, Freen PJW, Hoekman J, Maaskant MA. Body Mass Index (BMI) bij mensen met een verstandelijke beperking. Nederlands Tijdschrift voor de Zorg aan mensen met verstandelijke beperkingen. 2005;1:3-17.

6 Frey GC & Chow B. Relationship between BMI, physical fitness, and motor skills in youth with mild intellectual disabilities. Int J Obes. 2006;30:861-867.

7 Lahtinen U, Rintala P, Malin A. Physical performance of individuals with intellectual disability: a 30 year follow up. Adapt Phys Act Q. 2007;24: 125-143.

8 Draheim CC, Williams DP, MCCubbin JA. Prevalence of physical inactivity and recommended physical activity in community-based adults with mental retardation. Ment Retard. 2002;40:436-444.

9 Temple VA, Frey GC, Stanish HI. Physical activity of adults with mental retardation: review and research needs. Am J Health Prom. 2006;21:2-12.

10 World Health Organization. Physical status: The use and interpretation of anthropometry. 1995, Geneva, Switzerland: World Health Organization. (WHO Technical Report Series).

11 Deurenberg-Yap M, Schmidt G, Staveren WA van, Deurenberg P. The paradox of low body mass index and high body fat percentage among Chinese, Malays and Indians in Singapore. Int J Obes. 2000;24(8):1011-1017.

12 Gallagher D, Visser M, Sepúlveda D, Pierson RN, Harris T, Heymsfield SB. How useful is BMI for comparison of body fatness across age, sex and ethnic groups? Am J Epidem. 1996;143:228–239.

13 Hark L, Bowman M, Bellini L. Oversight of Nutrition in Clinical Care. In: Medical Nutrition & Disease (eds L. Hark & G. Morrison). 2003, National Health and Nutrition Examination Survey IV.

14 Dalton M, Cameron AJ, Zimmet PZ, Shaw JE, Jolley D, Dunstan DW, Welborn TA, Waist circumference, waist-hip ratio and body mass index and their correlation with cardiovascular disease risk factors in Australian adults. J Int Med. 2003;254(6):555-63.

15 Savva SC, Tornaritis M, Savva ME, Kourides Y, Panagi A, Silikiotou N, Georgiou C, Kafatos A. Waist circumference and waist-to-height ratio are better predictors of cardiovascular disease risk factors in children than body mass index. Int J Obes Rel Metab Dis. 2000;24(11):1453-8.

16 Mei Z, Grummer-Strawn LM, Pietrobelli A, Goulding A, Goran MI, Dietz WH. Validity of body mass index compared with other body-composition screening indexes for the assessment of body fatness in children and adolescents. Am J Clinic Nutr. 2002;7597–985.

17 Nadas J, Putz Z, Kolev G, Nagy S, Jermendy G. Intraobserver and interobserver variability of measuring waist circumference. Med Sci Mon. 2008;14(1):CR 15-18.
18 Nooyens AC, Koppes LL, Visscher TL, Twisk JW, Kemper HC, Schuit AJ, van Mechelen W, Seidell JC. Adolescent skinfold thickness is a better predictor of high body fatness in adults than is body mass index: the Amsterdam Growth and Health Longitudinal Study. Am J Clinic Nutr. 2007;85(6):1533-9.
19 Pischon T, Boeing H, Hoffmann K, Bergmann M, Schulze MB, Overvad K, et al. General and Abdominal Adiposity and Risk of Death in Europe. New Engl J Med. 2008;20:2105-2120.
20 Waninge A. Van der Weide W, Evenhuis, IJ, Van Wijck R,Van der Schans CP. Feasibility and reliability of body composition measurements in adults with severe intellectual and sensory disabilities. J Intellect Disabil Res. 2009;53:377-88.
21 Pollers S, Bayate M, Miedema-Loo HH, Evenhuis HM, Penning C. Overgewicht bij volwassen mensen met een lichte of matige verstandelijke handicap wonend in een instelling: het stellen van de diagnose en meetmethoden, 2007. Onderzoek in het kader van de opleiding tot AVG, Erasmus Universiteit Rotterdam.
22 Booth ML, HunterC, Gore CJ, Bauman A, Owen N. The relationship between body mass index and waist circumference: Implications for estimates of the population prevalence of overweight. Int J Obes Rel Metabol Dis. 2000;24:1058-1061.
23 Rosenbaum P, Paneth N, Leviton A. A report: the definition and classification of cerebral palsy. Dev Med Child Neur. 2007;109:8-14.
24 Bland JM & Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986;8:307-10.
25 Palisano R, Hanna SE, Rosenbaum PL, Rusell DJ, Walter SD, Wood EP, Raina PS, Galuppi BE. Validation of a model of Gross Motor Function for Children With Cerebral Palsy. Phys Therapy. 2000;80(10):974-985.
26 Dutch Society for Doctors in the Care for people with an Intellectual Disability (NVAZ). Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans, behavioural code for doctors in the assessment of resistance among people with an intellectual disability, 1999.
27 Bhaumik S, Watson JM, Thorp CF, Tyrer F, Mc Grother CW. Body mass index in adults with intellectual disability: distribution, association and service implications: a population-based prevalence study. J Intellect Disabil Res. 2008;52(4):287-98.
28 Emerson E. Underweight, obesity and exercise among adults with intellectual disabilities in supported accommodation in Northern England. J Int Dis Res. 2005;49(2):134-43.
29 Melville CA, Cooper SA, McGrother CW, Thorp CF, Collacott R. Obesity in adults with Down syndrome: a case-control study. J Intellect Disabil Res. 2005;49(2):125-33.
30 Prince SA, Janssen I, Tranmer JE. Self-Measured Waist Circumference in Older Patients With Heart Failure: A Study of Validity and Reliability Using a Myo Tape(R), J Cardiopulm Rehabil Prevention. 2008;28(1):43-47.
31 Kong CK & Wong HS. Weight-for-height values and limb anthropometric composition of tube-fed children with quadriplegic cerebral palsy. Pediatrics. 2005;116, 839-845.
32 Sullivan PB, Alder N, Bachlet AM Grant H, Juszczak E, Henry J, Vernon-Roberts A, Warner J, Wells J. Gastrostomy feeding in cerebral palsy: Too much of a good thing. Dev Med Child Neurol. 2006;48:877-82.

Chapter 4

Feasibility and reliability of two different walking tests in subjects with severe intellectual and sensory disabilities.

A. Waninge
I. J. Evenhuis
R. van Wijck
C.P. van der Schans

Journal of Applied Research in Intellectual Disabilities
Accepted for publication, July 2010

Abstract

Background. The purpose of this study is to describe feasibility and test-retest reliability of the six-minute walking distance test (6MWD) and an adapted shuttle run test (aSRT) in persons with severe intellectual and sensory (multiple) disabilities (SIMD).

Materials and Methods. Forty-seven persons with SIMD, with Gross Motor Function Classification System (GMFCS) grade I and II and wearing a heart rate monitor, performed the 6MWD and the aSRT twice.

Results. 96% Of the participants completed both tests successfully. Wilcoxon signed rank test revealed no significant differences between test and retest ($p<0.05$). Intra Class Correlationcoefficients for all variables were ≥.90. Limits Of Agreement for aSRT in GMFCS-II subjects were insufficient.

Conclusion. 6MWD is feasible and reliable for measuring functional exercise capacity in GMFCS-I and II participants with SIMD. aSRT is feasible and reliable for measuring aerobic capacity in GMFCS-I participants. Compared to others, participants with SIMD achieved poor results in 6MWD.

Introduction

People with intellectual disabilities (ID) make up about 1% of the population of Europe [1]. This percentage is based on the WHO population prevalence estimate [1]. Comorbidity in persons with intellectual disabilities is more frequent and patterns of comorbidity differ from the general population [2]. Obesity in women and underweight in both men and women is more common in adults with ID than in the general population after controlling for differences in the age distributions between the two populations [3]. Mc Guire et al. [4] found that 68% of their ID sample was overweight or obese, participation in exercise and adherence to a healthy diet are poor. Other authors described these lifestyle problems in adults with low or moderate ID too: these persons often suffer from overweight and may have poor physical fitness [5, 6, 7]. Adults with ID often are not sufficiently active to achieve health benefits [8, 9].

According to the Toronto model [10], physical fitness and health are related in the sense that good physical fitness decreases health risks and improves wellbeing and quality of life [11, 12, 13]. Health can be defined as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity (World Health Organization,WHO) [14, 15]. In addition, health is considered a resource for everyday life, not the objective of living. Health is a positive concept emphasizing social and personal resources, as well as physical capacities [16]. Bouchard et al. [17] give the following definition of health related physical fitness: 'Health related physical fitness is defined as a set of attributes that people have or achieve that relates to the ability to perform physical activity'. According to Caspersen et al. [18] and Pate [19], physical fitness is a set of attributes that are either health-related or skill-related that pertain more to athletic ability. The health-related components of physical fitness are (a) cardiorespiratory endurance, (b) muscular endurance, (c) muscular strength, (d) body composition, and (e) flexibility. These components of physical fitness are more important to public health than are the components related to athletic ability, especially in persons with multiple disabilities, therefore, we limit our discussion to these.

Wuang et al. [20]) described that the IQ level substantially predicted overall performance on motor tests. Physical fitness in persons with a visual disability is poorer than in persons without disabilities [21, 22, 23]. Furthermore, a considerable number of persons with both severe intellectual and sensory disabilities are at risk for a variety of health problems [24], among others inactivity, overweight and obesity. The women appear at a higher risk for developing health problems compared to the men [25]. Apart from that, these health burdens often are associated with low levels of physical fitness [8]. Therefore, it is important to get insight in the physical fitness status in these individuals.

Cardiorespiratory endurance, a component of health-related physical fitness [18], can be divided into aerobic power and cardiovascular capacity [17]. This implies that assessment of health-related physical fitness includes measures of aerobic capacity as well as functional exercise capacity [17], both objective measures of fitness level [26]. Aerobic capacity is the ability of the cardiovascular system to deliver oxygen rich blood to body tissues [17], functional exercise capacity is an objective measure of one's ability to undertake the activities of day-to-day life [27].

Timed walking tests are accepted methods to assess both above mentioned components of health-related physical fitness and to examine fitness level. Incremental speed walking tests (ISWT) are effective measures of aerobic capacity in healthy individuals and in individuals with chronic health conditions [28, 29]. These tests require participants to walk or run between two markers that delineate a 10-m course. Participants are to walk or run the course at a set

incremental speed determined by a signal, which is played by a standard CD player. Two newly developed shuttle run tests (SRT I, -II) for children with cerebral palsy also yield reliable and valid data for measuring aerobic power [30]. In these SRTs, speed increases every minute at 0.25 km/h steps. The outcome measure is the number of steps successfully completed at the time the test is stopped [30].

The six-minute walking distance test (6MWD) is accepted as a reliable test to measure functional exercise capacity in participants in various disease states [31, 32, 33, 34]. The 6MWD is self-paced and requires an individual to walk as far as possible in six minutes on a course of various lengths, without running. The distance walked during the test, measured in feet or metres, is used as the outcome measure [35].

The target population of this study are persons with severe or profound intellectual and sensory disabilities (SIMD). According to the ICD-10 (World Health Organization, WHO) [36], the IQ of persons with severe intellectual disabilities ranges from approximately 20 to 34, in adults this means an intellectual age from 3 to under 6 years, which is likely to result in continuous need for support. The IQ of persons with profound intellectual disabilities is below 20, in adults this means an intellectual age below 3 years, which results in severe limitation in self-care, continence, communication and mobility.

In particular persons with the combination of severe ID and sensory disabilities are at risk for both decrease of independence and quality of life, due to reinforcement of disabilities and less opportunities for compensation [37]. In these persons the level of health-related physical fitness is difficult to reliably quantify, because they are not accustomed to the above-mentioned assessments, such as performing walking tests. To assess whether a test is feasible for these persons, it is therefore necessary to use specific inclusion criteria to each of the co-morbidities e.g. locomotor disabilities and visual impairments. Furthermore, adjustments to test procedures are necessary, as well as choosing the tests carefully. For example, the SRT-I and II [30] may be most feasible for these persons, because of its increase of speed with 0.25 km/h, instead of the 0.6 km/h increase of speed of the ISWT [28, 29]. Another limiting factor to determine the feasibility and reliability of the measures in these persons are motivational problems. For instance, some do not understand why they have to walk faster than they usually do, or why they have to wear a heart rate monitor.

The feasibility and reliability of timed walking tests, such as the 6MWD and SRT, in participants with SIMD have been unknown. Because good physical fitness decreases health risks and improves wellbeing and quality of life [10, 11, 12], it is important to get comprehensive insight in the health related physical fitness in persons with SIMD. With feasible and reliable tests, a specific training intervention aimed at promoting physical fitness can be evaluated. Therefore, the purpose of this study was to examine the feasibility and the test-retest reliability of the 6MWD and SRT-I and II in persons with SIMD.

Method

Participants

The participants were recruited from a residential care facility in the Netherlands, in which 200 persons with severe or profound intellectual and sensory disabilities live. Moreover, in 65% they also exhibit associated motor disabilities. We asked the representatives of 92 persons with sufficient motor capacities for permission for the persons to participate in this study.

Eighty representatives gave permission. After informed consent was obtained, we screened these participants based on the examination findings of a physician specialised in intellectual disabilities and of a behaviour scholar and excluded seven participants. Another eight participants were excluded as they did not live at the centre for people with severe intellectual and sensory disabilities where the tests were performed. Eighteen participants were excluded because they presented with exclusion criteria at the time the tests were administered (Figure 1).

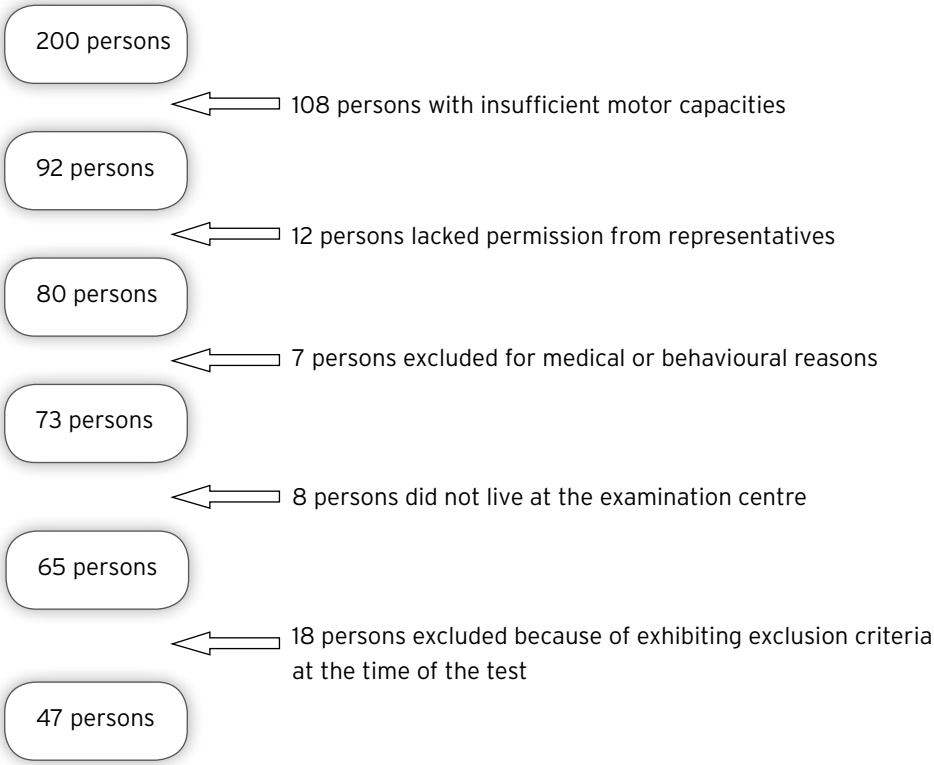


Figure 1 Inclusion steps

In all, 47 persons participated in this study: 18 were female, mean (SD) age 44 (10) years and 29 were male, mean (SD) age 38 (11) years. Participants were classified according to the adapted Gross Motor Function Classification System (GMFCS) [38]. Twenty-seven participants were classified as GMFCS level I, and 20 participants as GMFCS level II (see below for more information about GMFCS). Eighty-seven percent (41) of the participants had severe intellectual disabilities and 13% (6) had profound intellectual disabilities, according to the classification of the ICD-10 [36]. Most of the participants also had impaired vision. According to WHO guidelines [39], 53% (25) of the participants were severely partially sighted, 40% (19) were partially sighted, and 7% (3) were slightly limited in sight. Most participants had impaired motor abilities: 60% (28) had orthopaedic defects and 8% (4) were diagnosed with spasticity. In addition, 28% (13) of the participants had slight hearing problems, 9% (4) had loss of hearing, and 4% (2) had severe loss of hearing or were completely deaf.

As a group, individuals with severe or profound intellectual disabilities possess several co-morbidities simultaneously, and thus combinations of co-morbidities are also present in our

participants (Table 1). Thirty-six percent (17) of the intellectual disabled participants had both hearing problems and visual disabilities; 60% (28) had both visual and orthopaedic disabilities; 19% (9) had hearing problems and orthopaedic disabilities. 4 Participants were diagnosed with spasticity and they also had visual, and orthopaedic disabilities.

Table 1. Combinations of co-morbidities of participants with severe intellectual disabilities

	Hearing disabilities	Orthopaedic defects	Spasticity
Visual disabilities	36%	60%	4%
Hearing disabilities	-	19%	2%
Orthopaedic defects	-	-	9%

Participants were classified according to an adapted GMFCS [38], a five-level system used to classify the severity of motor abilities in people with physical disabilities. Participants with a "Level I" classification can generally walk without restrictions but tend to have limitations in some more advanced motor skills. Participants with a "Level V" classification generally have very limited mobility, even with the use of assistive technology. These participants always use a wheelchair. The original GMFCS was adapted because most of our participants had impaired vision, and as a result they could not jump and run spontaneously. If persons spontaneously increased their speed during walking, instead of jumping and running, they were classified as GMFCS level I. Participants with a "Level II" classification can walk with slight restrictions and do not spontaneously increase their speed during walking. The adapted version of the GMFCS was presented to the investigator, who translated the original version of the GMFCS into Dutch [40] and he concluded that the adaptations did not influence the reliability of the system.

Study Design

Participants were tested twice, with one week between the test and the retest. Test and retest were conducted at the same point in time. The participants performed first the aSRT (Netchild, the Netherlands). In order to let the participants take sufficient rest, at least after 48 hours, the 6MWD [35] was performed.

Ethical statement

The study was performed in agreement with the guidelines of the Helsinki Declaration as revised in 1975. Permission to carry out the study was obtained from an institutional ethics committee. All participants were unable to give consent. Therefore, extra care and attention was given to:
1) Asking informed consent: Informed consent was obtained from legal representatives of all participants and also the caregivers of all participants were asked for informed consent;
2) The construction of the study group by formulating exclusion criteria and contraindications: We screened the participants based on the examination findings of a physician specialised in intellectual disabilities and also of a behaviour scholar;

3) The measurement procedure: The measurements were performed in accordance with the behavioural code section entitled 'Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans' [42]. Consistent distress or unhappiness was interpreted as a sign of lack of assent and further participation in the study was reconsidered.

Protocols

Before the tests took place, the testing leaders and personal guides of the participants completed a checklist that included all contraindications. Participants were excluded from the study if they exhibited any of the following exclusion criteria at the time of the measurements: psychoses, depression, or other severe psychological problems; or somatic diseases, which were defined as chronic diseases and/or diseases that do not resolve in the short term (e.g., osteoarthritis, osteoporosis, pneumonia, etc). Participants were also excluded for the following reasons: general illness or fever; taking antibiotics; worsening of asthma, epilepsy (recent insult or epileptic fits), fresh wound(s)/bruise(s), or other factors causing pain during movement; or stress due to the subject's behaviour just before the measurement date. To reduce learning effects, the participants practised twice before formal testing began.

Six-minute walking distance test

Participants performed the 6MWD test on a 36-m course (Figure 2), which was located in a gymnasium. The participants walked six minutes at a self-chosen pace and tried to cover as much distance as possible without running. Instructors accompanied all participants to help them find their way. This was necessary because of the participant's visual disabilities. The participants were encouraged in a standardized way. The total distance covered represented a participants level of functional exercise capacity.

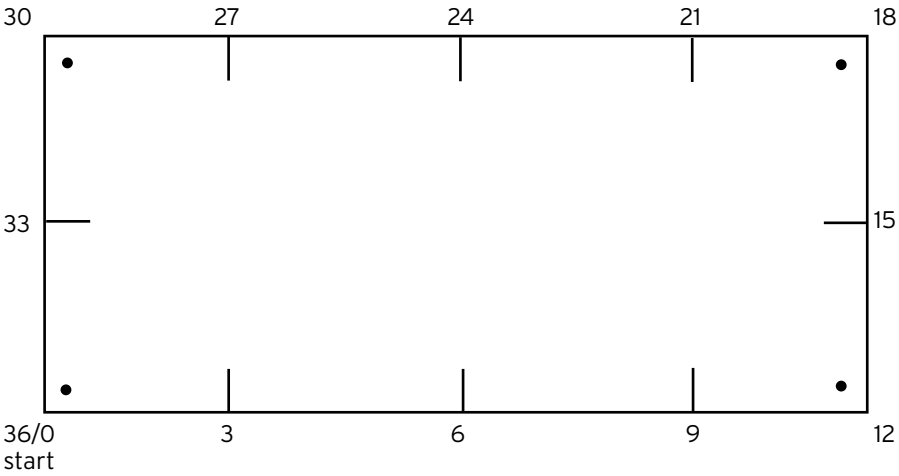


Figure 2 6MWD course

SRT-I, II and adapted Shuttle Run Test (aSRT)

Due to the severe multiple disabilities and the co-morbidities, practice sessions were performed in order to examine whether the protocols of the SRT I and II had to be adjusted. The SRT-I has a starting speed of 5 km/h, whereas the SRT-II has a starting speed of 2 km/h. During the abovementioned familiarisation period, we found that 5 km/h was too fast for our participants, but 2 km/h was too slow. Thus, we adjusted the starting speed to 3 km/h. In this adapted SRT (aSRT), speed was increased every minute at 0.25 km/h steps, conform the original procedure by Verschuren [30]. Every increase in speed is called a step. The number of steps successfully completed at the time the test is stopped, represents a participants level of aerobic capacity.

The aSRT course was located in a gymnasium and was composed of an oval curve with two markers (Figure 3). The subjects walked between two markers that delineated the 10-m course at a set incremental speed determined by an audio signal played from a standard CD player. Instructors accompanied all participants to help them pace themselves according to the audio signal. The course adaptations and the use of instructors were necessary because of the participants visual disabilities.

At the end of each step, the participants were told to walk a little faster. The test was finished when, on two consecutive paced signals, the participants were more than 1.5 m away from the end marker.

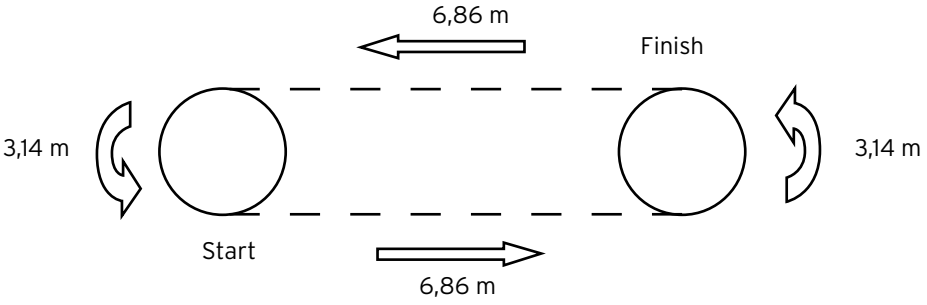


Figure 3 Shuttle test parcourse
R=1 Metre
Distance between Start and Finish = 6.86 m + ½ (2πR)= 6.86 + 3.14= 10 m

Motivation

To determine whether the participants pushed themselves to their limits and whether motivation influenced the tests, we used two strategies.

First, we compared the registered heart rate at the end of the test or when the test was stopped with the estimated peak heart rate. All participants wore a heart rate monitor (Polar Accurex plus™, Kempele, Finland), so that we could measure their heart rate every minute during performance and immediately after performance. During the tests, peak heart rate in beats per minute (bpm) was registered on a heart-rate-monitor storage device; these data could later be read from a wrist monitor and recorded on a datasheet. Each participant's peak heart rate was estimated using the formula of Fernhall [42] for participants with intellectual disabilities: 210 – 0.56 (age) – 15.5 (DS). DS is 1, however if a participant has Down's Syndrome, DS factor equals 2.

The second strategy to examine whether the participants pushed themselves to their limits

and whether motivation influenced the test results, was the use of the Visual Analogue Scale [43, 44]. The Visual Analogue Scale (VAS) is an instrument on which an instructor could score the degree of motivation on a 100-mm line with a stripe right-angled on the line. On one end of the line the maximum score was marked as 'good motivation' and on the other end was the minimum score, 'bad motivation'. The number of millimetres between the stripe scored by the instructor and the minimum score, is the score on the VAS instrument. Two instructors scored the degree of motivation.

Data analysis

Data were analysed using SPSS 14.0. Feasibility, test-retest reliability and motivation were examined as follows:

Feasibility

To assess feasibility, we compared the number of unsuccessful measurements with the total number of measurements. The feasibility was considered to be sufficient when 95% of the measurements were successful, which is strict, but only then it makes sense to use the test.

Test-retest reliability

First, to determine whether significant differences existed between the test and retest of the 6MWD, aSRT, and peak heart rate, we compared the test and retest data using Wilcoxon signed rank tests. The level of statistical significance was set at 0.05.

Limits of agreement (LOA) between the test and retest of the same variables were calculated according to the procedure described by Bland and Altman (1986). The LOA is considered to be an indicator of test – retest reliability. LOAs are expressed in units and as a percentage of the mean of the 1st test. Tests were considered to be reliable when the LOA was 30 % of the mean of the 1st test, or less, which is based on clinical experience of the professionals working with the target group and the measurements as suggested by Bland and Altman [45]. Finally, intraclass correlation coefficients (ICC, 2-way random, absolute agreement) of the test and retest of the same three variables were computed. Reliability was considered acceptable when the ICC value was greater than .80 and the 95% confidence intervals (CI) were 0.30 or less.

Motivation

To get insight into the influence of motivation on the test results, we first compared the registered heart rate at the end of both tests with the estimated peak heart rate using Wilcoxon signed rank tests. We also compared the registered heart rate at the end of the 6MWD with the registered heart rate at the end of the aSRT using Wilcoxon signed rank test. The level of statistical significance was set at 0.05.

Second, we calculated Spearman correlation coefficients for the number of aSRT steps achieved and motivation VAS score and for the 6MWD distance achieved and motivation VAS score. The level of statistical significance was set at 0.05. To calculate the influence of motivation on the test results we estimated the quadrate of the correlation coefficients and multiplied it with 100%. This is the variation which can be clarified with motivation [46]. The interrater reliability on the VAS for the aSRT and for the 6MWD were estimated with Spearman correlation coefficients, the level of statistical significance was set at 0.05.

Results

Feasibility

All 47 (100%) participants completed the tests successfully, and 45 (96%) of them also wore a heart rate monitor which produced valid data.

Test-retest reliability for all measurements

Table 2. summarizes the results of the Wilcoxon signed rank test, the LOA, the LOA as a percentage of the mean, and ICC analyses. There were no significant differences between the test and retest results in the 6MWD and aSRT. The LOAs and the LOAs expressed as a percentage of the means were less than or equal to 30% for all measurements, except for the aSRT of GMFCS-II participants. The ICCs were .80 or above. The LOA of the 6MWD was 115 m and that of the aSRT for GMFCS-I participants was 2.8 steps.

Table 2. Results of Wilcoxon ranking test, LOA, percentage LOA of mean, and ICC*

	Mean 1(SD)	**Mean 2 (SD)**	**P value Wilcoxon**	**LOA ***	**LOA of mean (%)**	**ICC* 95% CI**
SRT steps GMFCS I	12 (3.5)	12 (3.5)	0.422	2* 1.4023	23%	0.96 0.92-0.98
SRT steps GMFCS II	6 (3)	7 (3)	0.363	2* 2.4438	74%	0.82 0.54-0.93
6MWD distance GMFCS I and II	389 (104)	389 (109)	0.990	2* 57.7221	30%	0.92 0.86-0.96
Peak heart rate SRT	125 (20)	125 (19)	0.587	2* 14.348	22%	0.84 0.60-0.91
Peak heart rate 6MWD	117 (26)	119 (17)	0.907	2* 21.066	35%	0.84 0.71-0.91

*LOA, Limits of Agreement, ICC, Intraclass correlation coefficient, CI, Confidence Intervals

Motivation

Table 3. compares the mean peak heart rate achieved during both tests. The Wilcoxon signed rank test revealed a significant difference at $p<0.001$ between the registered heart rate at the end of both tests and the mean (SD) estimated peak heart rate (172 bpm ± 6). The Wilcoxon signed ranks test also identified a significant difference ($p<0.001$) between the estimated peak heart rate and peak heart rate achieved during the SRT (126 ± 20 bpm) and 6MWD (119 ± 16 bpm). The participants achieved a higher peak heart rate during the SRT than during the 6MWD.

Table 3. Mean peak heart rate achieved with aSRT and 6MWD*

	Mean peak heart rate (SD)	Mean estimated peak heart rate (SD)	p Wilcoxon test
aSRT	126 bpm (20)	172 bpm (6)	P<0.001
6MWD	119 bpm (16)	172 bpm (6)	P<0.001
p Wilcoxon test	p=0.01	-	-

*aSRT, adapted shuttle run test; 6MWD, six-minute walking distance test.

The Spearman correlation coefficient of SRT steps and motivation was 0.538 (p<0.01), and this correlation coefficient of 6MWD distance and motivation was 0.556 (p<0.01). The influence of motivation on the test results was calculated by quadrating the correlation coefficients and multiplying it with 100%. Thus, 29% of the variation we observed in the SRT step and 31% of the variation we observed in the 6MWD outcomes can be explained by motivation. The interrater reliability on the VAS for the aSRT was 0.87 (Spearman's rho; p<0.01) and for the 6MWD 0.92 (Spearman's rho; p<0.01).

Discussion

As far as we know, this is the first study in which the feasibility and test retest reliability in participants with severe intellectual and multiple disabilities is examined. The results of our study show that both the 6MWD and the aSRT, combined with heart rate monitoring, are feasible in participants with severe intellectual and sensory disabilities. Both tests can be reliably performed by GMFCS-I participants; GMFCS-II participants can only reliably perform the 6MWD.

The results are in line with the original procedures of the aSRT [30] and the 6MWD [33]: the aSRT appeared to be suitable for assessing aerobic capacity, because a higher peak heart rate was achieved with the aSRT, and the 6MWD is suitable for assessing functional exercise capacity.

Motivation is a term to express that someone tries to achieve a certain goal. We asked ourselves if persons with SIMD are motivated to exert themselves fully and if we get a realistic representation of what they really are able to. To determine whether the participants pushed themselves to their limits and whether motivation influenced the tests, we used the described two strategies: first, comparing the registered heart rate with the estimated peak heart rate and second, evaluating with the instructor the motivation of the participants on a VAS. We found that a participant's motivational level substantially can influence test outcomes. Twenty-nine percent of the variation in the SRT steps could be explained by motivation, whereas thirty-one percent of the variation in the 6MWD could be explained by motivation. Comparison of the estimated peak heart rate with the achieved peak heart rate shows that the participants did not fully exert themselves as they were performing the tests. We tried to motivate the participants by creating as optimal circumstances as possible: testing at the regular gymnastic time, in the regular gymnastic environment, accompanied by the familiar gymnastic instructors, who are used to motivate the participants. Our hypothesis is that the participants are not yet used to experience

the effects of considerable physical effort, due to their intellectual disabilities and reinforced by their visual disabilities. In order to get an adequate interpretation of the test results, we recommend to include both the comparison of estimated peak heart rate with the achieved peak heart rate and the score on motivation into the protocols of the aSRT and 6MWD.

Another complex question in these participants is informed consent: they are unable to give consent themselves. Therefore, extra care and attention was given to asking informed consent, the measurement procedure and the construction of the study group by formulating exclusion criteria and contraindications. All persons involved in the study were fully aware of the vulnerable position of the participants. If a participant did not like the testing, further participation was stopped.

In our study, the reliability of the 6MWD showed to be in line to that of similar measurements in other studies, and so was the aSRT. Several studies have used the 6MWD in participants having a variety of diseases. For participants with fibromyalgia, the ICC of 6MWD test-retest measurements was 0.73 [47]. For participants with heart failure, the ICC was 0.91 [48] and 0.82 [49]; and for participants with pulmonary disease, the ICC was 0.99 [50]. In our study, the ICC was 0.92, which is notable, because the comorbidities of the participants under study could potentially affect reliability. The reliability of the SRT-I and II was described by Verschuren et al.[30], who obtained an ICC of 0.97 for GMFCS-I participants. In our study, the ICC for comparable participants was 0.96. For GMFCS-II participants, Verschuren et al. [30] obtained an ICC of 0.99. In our study, the ICC was 0.82; however, the 95% CI were too wide and the LOA of the aSRT was unacceptable for GMFCS-II participants. This suggests, that aerobic capacity cannot be reliably assessed in GMFCS-II participants. A hypotheses might be that their locomotor skills do not allow them to increase their speed consequently.

The participants in our study first practiced twice before we measured their performance on the 6WMD and aSRT. Stevens et al. [51] described several studies that showed that practise sessions are necessary to promote optimal performance in participants. These investigators indicated that two practise sessions are required in order to allow participants to learn and to establish optimal performance when conducting repeated measures at relatively short intervals. This was also necessary for the participants of our study.

We compared the achieved mean distance in the 6MWD of our participants with that of others. The mean distance (SD) of our participants was 389 m (107). In healthy elderly persons, the mean distance was 631 m (93) [33]; in people with heart failure, the mean distance was 419 m (120) [31]; and in people with COPD, the mean distance was 369 m (18) [32]. This comparison indicates that persons with severe multiple disabilities performed poorer on the 6MWD than other persons with specific (chronic) health conditions.

In conclusion, the 6MWD is feasible and reliable for measuring functional exercise capacity in GMFCS-I and II participants with severe intellectual and multiple disabilities. The aSRT is feasible and reliable for measuring aerobic capacity in GMFCS-I participants. The participant's motivational level can influence test outcomes, so we recommend to insert both heart rate monitoring and motivational score into the protocols of the aSRT and 6MWD. The poor 6MWD results we observed indicate that the poor functional exercise capacity of persons with severe multiple disabilities is a serious health problem, which may burden their independence in day-to-day activities. Based on this result, further research should be aimed at developing, implementing and evaluation of an appropriate intervention to reduce problems in functional exercise and aerobic capacity in these participants.

Acknowledgements

This research was financed by Hanze University Groningen, De Brink, and with funding from the Regional Action-and-Attention Knowledge Circulation (RAAK).

The authors kindly acknowledge and thank the participants for their participation in this study, their representatives for giving permission to this and the gymnastic instructors of The Brink, for accompanying the participants during the walking tests.

Heart rate was registered by a heart rate monitor of Polar Accurex plus™, Kempele, Finland. The protocol of the Shuttle Run Test was provided by Netchild, the Netherlands. For the 6 MWD the protocol of Butland et al (1982) was used.

References

1 Walsh PN, Kerr M,Van Schrojenstein Lantman-de Valk HM. Health indicators for people with intellectual disabilities: a European perspective. Eur J Public Health. 2003;13(3 Suppl):47-50

2 Van Schrojenstein Lantman-de Valk HMJ, Metsemakers JFM, Haveman MJ, Crebolder HFJM. Health problems in people with intellectual disability in general practice: A comparative study. Family Practice. 2000;17:405–407

3 Bhaumik S, Watson JM, Thorp CF, Tyrer F, Mc Grother CW. Body mass index in adults with intellectual disability: distribution, association and service implications: a population-based prevalence study, J Intellect Disabil Res. 2008;52(4):287-98.

4 McGuire BE, Daly P, Smyth F. Lifestyle and health behaviours of adults with an intellectual disability. J Intellect Disabil Res. 2007;51(7):497-510.

5 Frey GC & Chow B. Relationship between BMI, physical fitness, and motor skills in youth with mild intellectual disabilities. Int J Obes. 2006;30:861-867.

6 Knijff-Raeven van AGM, Jansen-Jacobs CCM, Freen PJW, Hoekman J, Maaskant MA. Body Mass Index (BMI) bij mensen met een verstandelijke beperking. Nederlands Tijdschrift voor de Zorg aan mensen met verstandelijke beperkingen. 2005;1:3-17.

7 Lahtinen U, Rintala P, Malin A. Physical performance of individuals with intellectual disability: a 30 year follow up. Adapt Phys Act Q. 2007;24: 125-143.

8 Draheim CC, Williams DP, MCCubbin JA. Prevalence of physical inactivity and recommended physical activity in community-based adults with mental retardation. Ment Retard. 2002;40:436-444.

9 Temple VA, Frey GC, Stanish HI. Physical activity of adults with mental retardation: review and research needs. Am J Health Prom. 2006;21:2-12.

10 Bouchard C, Shepard RJ, Stephens T. Physical activity, Fitness and Health. 1994, Human Kinetics Publishers, ChampaignIL.

11 Colcombe S, Erickson KI, Raz N, Webb AG, MC Auley EB, Kramer AF. Aerobic fitness reduces brain tissue loss in aging humans. J Gerontol. 2003;58:176-180.

12 Kramer AF, Hahn S, Cohen NJ, Banich MT, McAyley E, Harrison CR, Chason J, Vakil E, Bardell L, Boileau RA, Colcombe A. Ageing, fitness and neurocognitive function. Nature. 1999;400:418-419

13 Sitskoorn, MM. Het plastische brein; invloed van gedrag. De Psycholoog. 2005;40(5): 262-267.

14 World Health Organization, Health for all, 1979.

15 World Health Organization, WHO Chr, 1980;34(2):80.

16 World Health Organization. Ottawa Charter for Health Promotion, 1986.

17 Bouchard C, Shephard RJ, Stephens T.(Eds.). Physical Activity, Fitness and Health; International proceedings and consensus statement. Human Kinetics Publishers, Chaimpaign IL. 1994;12:172-174.

18 Capsersen CJ, Powell KE, Christenson GM. Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research. Public Health Reports. 1985;100(2):127-131.

19 Pate RR, The evolving definition of fitness. Quest. 1988;40: 174-179.

20 Wuang YP, Wang CC, Huang MH, Su CY. Profiles and cognitive predictors of motor functions among early school-age children with mild intellectual disabilities. J Intellect Disabil Res. 2008;52(12):1048-60.

21 Hopkins WG, Gaeta H, Thomas AC, Hill PM. Physical fitness of blind and sighted children. Eur J of Appl Phys Occ Phys. 1987;56: 69-73.
22 Häkkinen A, Holopainen E, Kautiainen H, Sillanpää E, Häkkinen K. Neuromuscular function and balance of prepubertal and pubertal blind and sighted boys. Acta Paediatrica 2006;95(10):1277-83.
23 Seemungal BM, Glasauwer S, Gresty MA, Bronstein AM. Vestibular perception and navigation in the congenitally blind. J of Neurophys. 2007;97(6):4341-56.
24 Evenhuis HM, Sjoukes L, Koot HM, Kooijman AC. Does visual impairment lead to additional disability in adults with intellectual disabilities? J Intellect Disabil Res. 2009;53(1):19-28.
25 Waninge A. Van der Weide W, Evenhuis, IJ, Van Wijck R, Van der Schans CP. Feasibility and reliability of body composition measurements in adults with severe intellectual and sensory disabilities. J Intellect Disabil Res. 2009;53:377-88.
26 Fernhall B, Millar AL, Tymeson GT, Burkett LN. Maximal exercise testing of mentally retarded adolescents and adults: reliability study. Arch Phys Med Rehabil. 1990;71:1065-1068.
27 Guyatt GH, Sullivan MJ, Thompson PJ, Fallen EL, Pugsley SO, Taylor D.W. The six-minute walk: a new measure of exercise capacity in patients with chronic heart failure. Can Med Ass Journal. 1995;132:919-923.
28 Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of shuttle walking test of disability in patients with chronic airways obstruction. Thorax. 1992;47:1019-1024.
29 Singh SJ, Morgan MD, Hardman AE, Rowe C, Bardsley PA. Comparison of oxygen uptake during a conventional treadmill test and the shuttle walking test in chronic airflow limitation. Eur Resp Journal. 1994;7:2016-2020.
30 Verschuren O, Takken T, Ketelaar M, Gorter JW, Helders PJM. Reliability and validity of data for 2 newly developed shuttle run tests in children with cerebral palsy. Physical Therapy. 2006;86:1107-1117.
31 Faggiano P, D'Aloia A, Gualena A, Lavatelli A, Giordano A. Assessment of oxygen uptake during the 6-minute walking test in patients with heart failure: preliminary experience with a portable device. Am Heart Journal.1997;134: 203-206.
32 Onorati P, Antonucci R, Valli G, Berton E, De Marco F, Serra P. Non-invasive evaluation of gas exchange during a shuttle walking test vs. a 6-min walking test to assess exercise tolerance in COPD patients. Eur J Appl Phys. 2003;89:331-336.
33 Troosters T, Gosselink R, Decramer M. Six minute walking distance in healthy elderly subjects. Eur Resp Journal. 1999;14:270-274.
34 Troosters T, Vilaro J, Rabinovich R, Casas A, Barbera JA, Rodriguez-Roisin R. Physiological responses to the 6-min walk test in patient with chronic obstructive pulmonary disease. Eur Resp Journal. 2002;20:564-569.
35 Butland RJA, Pang J, Gross ER, Woodcock AA, Geddes DM. Two-, six- and 12-minute walking test in respiratory disease. Brit Med Journal. 1982;284:1607-1608.
36 World Health Organization. International statistical classification of diseases and related health problems: tenth revision (ICD-10). 1992;369–370.
37 Kiestra T, De unieke handicap, referentiemodel voor meervoudige beperkingen. 2005, ISBN 9080638420, Scholma Druk, Bedum
38 Palisano R, Hanna SE, Rosenbaum PL, Rusell DJ, Walter SD, Wood EP, Raina PS, Galuppi BE. Validation of a model of Gross Motor Function for Children With Cerebral Palsy. Physical Therapy. 2000;80:974-985.

39 World Health Organization (WHO). International Classification of Functioning, Disability and Health. Geneva, 2001
40 Gorter JW. Gross Motor Function Classification System (Dutch translation). 2001, Utrecht: Revalidatiecentrum De Hoogstraat.
41 Dutch Society for Doctors in the Care for people with an Intellectual Disability (NVAZ). Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans, behavioural code for doctors in the assessment of resistance among people with an intellectual disability, 1999.
42 Fernhall B, McCubbin J A, Pitetti KH, Rintala P, Rimmer JH, Millar AL, De Silva A. Prediction of maximal heart rate in individuals with mental retardation. Med Sci Sports Exerc. 2001;33:1655-1660.
43 Mc Cormack HM, Horne de DJL, Sheather S. Clinical applications of visual analogue scales: a critical review. Psych Med, 1988;18:1007-1019.
44 Wewers ME & Lowe NK. A critical Review of Visual Analogue Scales in the Measurement of Clinical Phenomena. Res Nursing Health. 1990;13:227-236.
45 Bland JM & Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986;8:307-10.
46 Baarda DB, Goede de MPM, Dijkom van CJ. Basisboek Statistiek met SPSS. Groningen/ Houten (The Netherlands): Wolters-Noordhoff bv; 2003:177.
47 King S, Wessel J, Bhambhani Y, Maikala R, Sholter D, Maksymowych W. Validity and reliability of the 6 minute walk in persons with fibromyalgia. J Rheumatol. 1999;26:2233-2237.
48 O'Keeffe ST, Lye, M, Donnellan C, Carmichael DN. Reproducibility and responsiveness of quality of life assessment and six minute walk test in elderly heart failure patients. Heart. 1998;80:377-382.
49 Roul G, Germain P, Bareiss P. Does the 6-minute walk test predict the prognosis in patients with NYHA class II or III chronic heart failure? Am Heart Journal. 1998;136:449-457.
50 Cahalin LP, Mathier MA, Semigran MJ, Dec GW, Disalvo TG. The six minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. Chest. 1996;110:325-332.
51 Stevens D, Elpern E, Sharma K, Szidon P, Ankin M, Kesten S. Comparison of hallway and treadmill six-minute walk tests. Am J Resp Critic Care Med. 1999;160:1540–1543.

Chapter 5

Psychometric quality of a graded treadmill exercise test for people with severe or profound intellectual and visual disabilities.

C.H. Sickler
A. Waninge
T. Takken
C.P. van der Schans
R. van Wijck

Submitted

Abstract

Introduction Exercise tests using treadmills are valuable tools for assessing exercise capacity. However, a treadmill protocol for persons with severe or profound intellectual and visual disabilities (severe or profound intellectual and multiple disabilities, SIMD) is not yet available.
Objective The present study investigated primarily the feasibility, validity and test-retest reliability of a graded treadmill exercise test (GXT) for people with SIMD and GMFCS level I.
Method Thirty participants with SIMD and GMFCS level I performed a graded exercise test and retest. A supra maximal block test (SMBT) was administered to assess validity. Two physical therapists alternated as test leaders, assisted by one of seven specially trained physical education instructors.
Results The participants' mean (sd) age was 41 years (11 yrs). Feasibility was sufficient for the test and retest of the GXT (86.6%). For the SMBT feasibility was less than sufficient (76.9%). Correlation of the peak heart rate (HR_{peak}) between the 1st GXT (GXT1) and 2nd GXT (GXT2) was good and significant (ICC=0.95; 95%CI 0.88-0.98) with good agreement (t-statistic p=0.5). Limits of agreement (LOA) were -16 to 14, which amounts to 19.5% of the mean HR_{peak}. The number of attained levels of the GXT1 and GXT2 showed a highly significant correlation (ICC=0.95; 95%CI 0.90-0.98). Correlation between HR_{peak} GXT and HR_{peak} SMBT was good (ICC=0.94; 95%CI 0.86-0.98) with good agreement (t-statistic p=0.7). LOA were from -17 to 15, being 20.5% of the mean HR_{peak}. Validity of GXT was good. As a secondary result, correlations and agreements between directly measured HR_{peak} and estimated HR_{peak} (using the Fernhall equation) were poor.
Conclusion A GXT performed on a treadmill is a feasible, reliable and valid means of determining HR_{peak} as well as number of attained levels for people with SIMD and GMFCS level I. At the individual level, results may have fairly large variability. The Fernhall equation for estimating HR_{peak} for people with SIMD systematically overestimated HR_{peak}.

Introduction

Intellectual disability (ID) is characterized by significant limitations in both intellectual functioning and adaptive behaviour as expressed in conceptual, social, and practical skills [1]. The disability originates before the age of 18 [1]. Intellectual disabilities are categorized in four groups: mild, moderate, severe and profound. Persons with severe or profound ID have a prevalence of visual impairments of 92% [2].

People with intellectual disabilities tend to have low activity and fitness levels which decline over the years when compared with those without a disability [3, 4]. Like individuals with intellectual disabilities, persons with visual impairments also display poor performance in locomotor skills [5] and have low levels of habitual activity [6]. Individuals who suffer from a combination of severe or profound intellectual and visual disabilities (severe or profound multiple disabilities, SIMD) are particularly at risk in terms of the potential development of deficits in both locomotor skills and daily functioning [7]. Furthermore, the combination of these deficits suggests that persons having SIMD are likely to have lower levels of habitual activity, than persons with ID without visual impairment.

Bouchard et al. [8] describes the relationship between physical activity, health-related fitness and health. Good physical fitness improves wellbeing and quality of life [9, 10], and decreases health risks, such as overweight and obesity [11]. Bouchard et al. described cardio-respiratory fitness as an important component of health-related fitness [8]. However, a considerable number of persons with SIMD achieved poor results compared to other persons with specific health conditions on health-related fitness as measured by the six minutes walking test [12]. These findings hence underscore the importance of gaining comprehensive insight into the health-related fitness of persons with SIMD, including the level of cardiorespiratory fitness.

Waninge et al. [12] performed a feasibility and reliability study for an adapted Shuttle Run Test (aSRT) in adults with SIMD. Feasibility and test-retest reliability of this aSRT over ground were good for participants classified on the Gross Motor Functional Classification Scale (GMFCS) as level I, yet not sufficient for those with GMFCS level II [12]. The peak heart rate (HR_{peak}) for each participant was estimated according to the Fernhall equation [13]. However, a significant difference ($p<0.001$) was found between the mean (±sd) HR_{peak} at the end of the aSRT (126 beats/min ± 20) over ground and the mean (±sd) estimated HR_{peak} (172 beats/min ± 6) [12]. This finding suggests that the aSRT overground has a limited validity as so far as the objective of the test is to measure the HR_{peak}. However, an alternative explanation of the significant difference between the achieved and the estimated HR_{peak} may point to the formula designed by Fernhall et al [13], which could simply overestimate the exercise capacity of subjects with SIMD.

Treadmill tests are known to obtain valid data for cardiorespiratory fitness [14, 15], and thus may help to solve this question. Testing a group of participants with SIMD using a graded treadmill test will give us an answer as to whether their peak heart rates are lower than expected, thereby rendering the Fernhall equation [13] invalid, or whether the aSRT test [12] is not able to yield a peak heart rate for this target group.

Another advantage of using a treadmill is that this apparatus might counteract the well-known tendency of persons with SIMD to display a lower motivation for physical performance than those without ID [3, 16]. Environmental cues as well as the gradually increasing speed of the treadmill reduce the degree of freedom as described in theories about perception-action coupling and the 'constraint-led approach' [17], thereby facilitating physical performance by our SIMD participants.

In order to confirm the validity of the aSRT [12], results of both the treadmill and aSR test should be compared to a golden standard [18]. Midgley et al. [19] reviewed available literature regarding the verification phase following a graded exercise test and suggested performing a Supra Maximal Block Test (SMBT) to set a gold standard for HR_{peak} on an individual basis [19]. If the heart rate of the SMBT differs by no more than two beats from the HR_{peak} reached during the corresponding aSRT [12], the exertion during the aSRT [12] is scored as maximal [19].

The purpose of this study thus is twofold; it seeks to determine the feasibility, validity and test-retest reliability of a graded treadmill test (GXT) for participants with SIMD and GMFCS level I, using a Supra Maximal Block Test (SMBT) as golden standard. By doing so, this study will also assess whether the formula of Fernhall overestimates HR_{peak} in persons with SIMD.

Methods

Participants

The participants were recruited from a residential care facility in the Netherlands, which is home to 200 persons with severe or profound intellectual and visual disabilities. Only persons functioning at GMFCS-I level were included, because Waninge et al. have indicated the aSRT performed over ground is only reliable for persons with GMFCS level I [12]. Probably due to visual impairments, the participants were not able to run and jump spontaneously [5, 7, 12]. Subsequently, classification was adjusted so as to include people with GMFCS-I who were able to increase walking speed [12].

A total of 30 participants were recruited (17 males, 13 females), all classified as GMFCS-I. Twenty-three participants had a severe ID and seven a profound ID. Some could not walk alone because of a visual impairment. Table 1 presents the participants' characteristics.

Table 1. Population

		Male (n=17)	**Female (n=13)**	**Total (n=30)**
Age	Mean (SD)	40 (11)	42 (11)	41 (11)
ID	Severe ID (20≤IQ<35) Profound ID (IQ<20)	12 5	12 2	23 7
Down Syndrome	Number	1	1	2
Vision	Blind Impaired vision	11 6	8 5	19 11
GMFCS	I	17	13	30

n: sample size; SD: Standard Deviation; ID: Intellectual Disability; GMFCS: Gross Motor Function Classification System; IQ: Intelligence Quotient

Exclusion criteria were mental or physical health problems that prevented the client from participating. A physician, specialised in the care for people with ID, approved participation. Exclusion criteria at the moment of testing were illness or fever, recent epileptic insult, stress, wound, concussion or pain during movement.

Design

Participants were tested twice, the first graded treadmill test (GXT1) was followed by a second graded treadmill test (GXT2) at least one week later. Both tests lasted 15-25 minutes. The same instructor and the same test leader performed the testing procedure at the same time of day. The test leader scored HR_{peak}, attained level and test time.
A two-minute supra-maximal block test (SMBT) followed at random either GXT1 or GXT2 (convenience sampling). A Trimline T370HR treadmill (Tunturi, Almere, the Netherlands) was used for testing. A Polar RS 800 heart rate monitor (Polar Nederland, Almere, the Netherlands) measured heart rate during the test.

Ethical statement

The study was performed in agreement with the guidelines of the Helsinki Declaration as revised in 1975. Permission to carry out the study was obtained from an institutional ethics committee. All participants were unable to give consent. Therefore extra care and attention was given to:
1) Obtaining informed consent. Informed consent was obtained from legal representatives and caregivers of all participants;
2) The construction of the study group by formulating exclusion criteria and contraindications: We screened the participants based on the examination findings of a physician specialised in intellectual disabilities and also of a behaviour scholar;
3) The measurement procedure: The measurements were performed in accordance with the behavioral code section entitled 'Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans' [20]. Consistent distress or unhappiness was interpreted as a sign of lack of assent and further participation in the study was reconsidered.

Protocols

Graded treadmill test (GXT)

The adjusted SRT protocol of Waninge et al. was used and performed on a treadmill [12]. All participants practiced walking the treadmill at least twice before being tested. The specially trained physical education instructor (instructor) accompanied the participant to the treadmill and attached the safety cord. The instructor explained the procedure and attached the belt of the heart rate monitor to the participant. To ensure safety the instructor positioned himself behind the participant, with a foot on each side of the treadmill. The test leader checked the heart rate monitor and stayed one meter on the side of the participant during the whole test procedure.

Rintala et al. reviewed the familiarization process in cardio-respiratory fitness testing in persons with mild to moderate intellectual disabilities, recommending a familiarization protocol [21]. To reduce stress due to unfamiliar situations, the participants walked the treadmill at least twice at regular walking speed before testing. Participants did not reach their HR_{peak} levels during these practice sessions.

The treadmill stood in a gymnasium where other clients were doing workouts. The test leader calibrated the treadmill once a week to guarantee reliability. The starting speed was 3 km/hour for all participants. Each minute the speed increased with alternating 0.3 and 0.2 km/hour increments to reach the next level. Levels were comparable with the levels in the Waninge study [12], in which 0.5 km/h increase of speed appeared not feasible for persons with SIMD.

The test continued until volitional exhaustion, similarly to the study of Rintala et al [16]. The test was finished in three possible ways; when the participant insisted he wanted to stop or refused to continue, when the predicted HR_{peak} [13] was reached, or when the safety pin was pulled. The level maintained during the last full minute stage counted as the highest attained level. After either GXT1 or GXT2, the participant rested for five minutes and went on to perform the SMBT.

Seven instructors participated in data collection. Two physical therapists acted as test leaders. They checked exclusion criteria at the moment of testing on a form filled out by the participant's personal care professional.

Supra Maximal Block Test (SMBT)

The individual gold standard for HR_{peak} was set by SMBT [19]. Midgley et al [19] reported the utility of the verification phase, which was performed at a speed equivalent to one stage higher than that attained during the last completed stage of the incremental phase. They did not exactly define the incremental phases and corresponding speed, as they found that 'despite the incremental phases being distinctly different, the mean maximal VO2max values attained in the appended verification phases were almost identical' [19]. Therefore, in the present study, speed during the verification phase was 0.2 or 0.3 km/h higher than speed during the last completed stage, as in persons with SIMD only an increase of speed of 0.2 or 0.3 km/h is feasible [12]. This is considered an acceptable increase of speed, because Midgley et al [19] stated that 'the verification phase should incorporate a workload higher than that attained in the incremental phase to conform to the original concept of VO2max'. If the heart rate of the SMBT differed by not more than two beats from the HR_{peak} reached during the corresponding GXT, the performance during the GXT was scored as maximal [19]. During the two minutes of the verification phase, the attained HRpeak was registered in beats per minute.

Two alternative protocols were developed. Criterion A was the SMBT [19]. The participant walked for another two minutes at a level one step (0.2 or 0.3 km/hr) higher than the level attained during the GXT. If the instructor thought, for instance for behavioural reasons, an increase in level was not feasible, the participant performed criterion B, which was a two minutes maximal block test at the highest previously attained level. The HR_{peak} was registered in beats per minute. Validity calculation did not include results from criterion B.

Motivation

Since people with ID tend to have lower motivation for physical activity, motivation was considered as a factor influencing the validity of the test [3, 16]. Paired modelling and positive reinforcement have a positive effect on compliance to treadmill walking for people with moderate to severe ID [22]. Participants were encouraged to continue walking using these techniques. Both the instructor and the test leader observed independently the amount of encouragement given as well as the compliance with the task. The observed motivation was defined as to how well the

participant had fulfilled the task. This was drawn out using a 100 mm Visual Analogue Scale (VAS) [23, 24]. Zero corresponded to no motivation and 100 to the best possible motivation.

Estimated peak heart rate

Several researchers have found that the HR_{peak} for people with ID systematically differs from the HR_{peak} of people without ID [25,26]. Therefore, Fernhall et al. developed an equation to more accurately predict HR_{peak} for people with ID [13]. The estimated HR_{peak} using the Fernhall formula [13] was calculated for each participant before testing.

Data Analysis

All statistical analyses were performed using the Statistical Package for Social Studies (SPSS) version 16.0 for Windows. All data was checked for normal distribution and homoscedasticity by statistical analysis. Homoscedasticity was defined as no relation between the error and the size of the measured value [27]. A p-value of < 0.05 indicated statistical significance for all tests.

Feasibility

The GXT and the SMBT were tested for feasibility. The percentage of the participants that finished the test successfully determined feasibility. Interpretation of the feasibility scores was taken from the Groningen fitness test for the elderly [28]. A 95% score meant "good" feasibility and an 80% score meant "sufficient" feasibility [28].

Test-retest reliability

The Intraclass Correlation Coefficient (ICC) (two-way random, absolute agreement) was calculated for HR_{peak} and for the attained GXT1 and GXT2 levels. An ICC<0.75 indicated poor or moderate reliability, 0.75≥ICC<0.90 good reliability and ICC0.90 very good reliability [29].

Agreement was analysed with a paired samples t-test for HR_{peak} and attained levels. In a Bland-Altman plot the Limits of Agreement (LOA) was determined [30]. The LOA expressed as a percentage of the mean described the variability of the results. The LOA represents individual variability. No criteria are available for 'good', 'sufficient' or 'poor' LOA. In literature researchers themselves are encouraged to judge whether the LOA is narrow enough for the test to be of practical use [27].

To assess test-retest reliability the Standard Error of Measurement (SEM), the Smallest Detectable Difference (SDD) and Effect Size (ES) were calculated for HR_{peak} and the attained GXT and GXT2 levels.

The Standard Error of Measurement (SEM) represents the standard deviation of measurement error. The SEM reflects the reliability of the response [29] and indicates the maximum difference that could be based on measurement error. SEM was calculated using the formula: $SEM=SD \times \sqrt{(1-ICC)}$ [27].

The Smallest Detectable Difference (SDD) reflects the diversity of the participants. It is a measure of agreement and is based on measurement error. The SDD represents the smallest true difference. The SDD was calculated as $SDD = 1.96 \times \sqrt{(2 \times SEM)}$ [27, 31].

Clinical relevance was estimated calculating the Effect Size (ES). The ES gives an objective and standardised measure for observed differences [32]. The formula for calculating ES: $r = \sqrt{\{t^2/(t^2+df)\}}$ [32], t being the t-statistic and df the degrees of freedom. The cut-off point for a small ES is 0.1, for a medium ES 0.3 and for a large ES 0.5 [32].

Validity

Validity calculation only includes results of the SMBT, performed according to criterion A (see Protocols). To evaluate validity we compared the HR_{peak} of the SMBT with the HR_{peak} of the corresponding GXT, which was either GXT1 or GXT2. The ICC was calculated for HR_{peak}. Agreement was analysed with a paired samples t-test.

In a Bland-Altman plot the LOA was determined [30]. The LOA expressed as a percentage of the mean described the variability of the results.

Furthermore, to assess validity the SEM, SDD and ES were calculated for the HR_{peak} of the SMBT and the HR_{peak} of the corresponding GXT.

The Spearman correlation was calculated for the motivation scored by the instructor and by the test leader.

The ICC (two-way random, absolute agreement) for HR_{peak} and level was used to compare the results of the aSRT over ground of Waninge et al. and the GXT on the treadmill in the same participants. Agreement was analysed with a Wilcoxon signed ranks test for both HR_{peak} and level.

Estimated peak heart rate

We analysed validity of the equation developed by Fernhall et al. in persons with SIMD, comparing measured HR_{peak} with the estimated HR_{peak} [13]. The HR_{peak} SMBT was used in a regression analysis to calculate β for each participant. To minimize the influence of outliers the five highest and five lowest heart rate scores of the SMBT were eliminated. The equation developed by Fernhall et al. was adjusted by a new β-constant useful for our sample.

Results

Feasibility

The test-retest feasibility of both GXTs was 86.6%, which can be considered as sufficient [28]. Twenty-six out of thirty participants completed both GXT1 and GXT2 successfully. One participant dropped out because of a heart rate irregularity during the first GXT. Two participants dropped out because the GXT1 caused stress and behavioural problems. One participant consistently stepped on the sides of the treadmill and was not able to perform the GXT.

The feasibility of the SMBT was 76.9%, which fails to be sufficient [28]. A total of twenty out of twenty-six participants who completed the GXT1 and GXT2 performed the SMBT successfully (criterion A). Two participants did not perform the SMBT because of a high stress level. Four persons performed the maximal block test according to criterion B.

Table 2. Descriptive results peak heart rate and levels

	HR_{peak} GXT1 (n=26)	HR_{peak} GXT2 (n=26)	HR_{peak} GXT (n=20)	HR_{peak} SMBT (n=20)	HR_{peak} calculated (n=26)	Level GXT1 (n=26)	Level GXT2 (n=26)
Mean (beats/min)	154	155	156	157	171	13.54	13.77
(SD)	(16)	(17)	(19)	(16)	(7)	(3.9)	(3.7)
Range (beats/min)	118-181	117-186	118-181	118-178	159-184	6-22	4-22

HR: heart rate (beats/min); GXT1: 1st GXT; GXT2: 2nd GXT; GXT: test corresponding to the SMBT; SMBT: supra-maximal block test; SD: standard deviation (beats/min)

Test-retest reliability

Table 2 represents the descriptive statistics of the HR_{peak} and the attained levels of GXT1, GXT2 and SMBT. The HR_{peak} data in GXT1 and GXT2 were normally distributed.

Pearson's correlation between the HR_{peak} GXT1 and the difference in HR_{peak} between GXT1 and GXT2 was low (r =0.11), which is indicative of the homoscedasticity of the results.

The ICC (ICC=0.95; 95%CI 0.88-0.98) for the HR_{peak} of the GXT1 and GXT2 was very good (Table 3). The t-test showed no significant difference between the measurements (p=0.5) (Table 3).

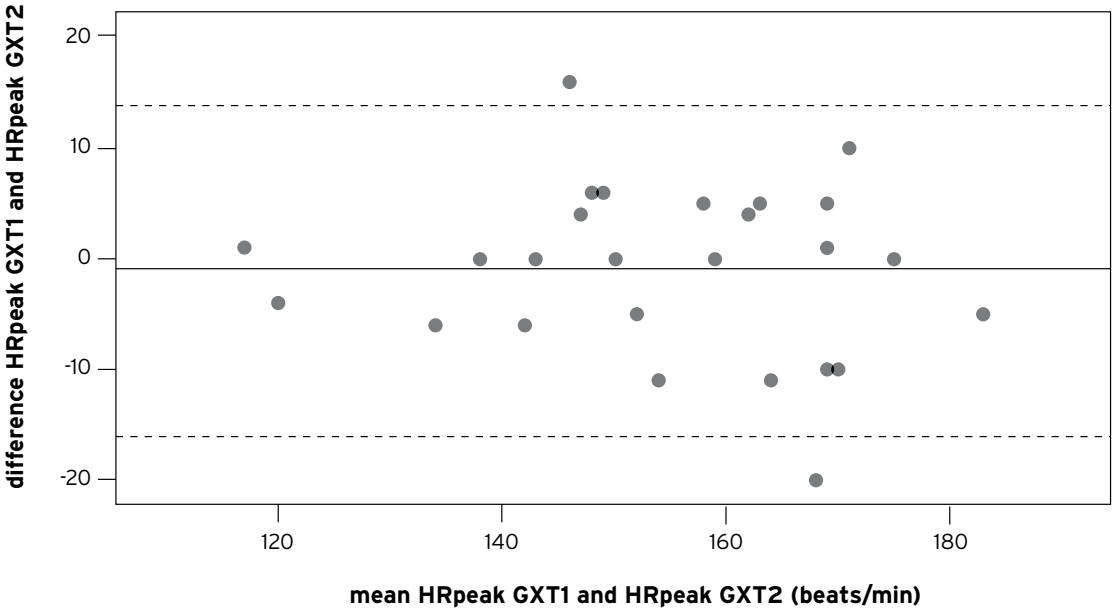


Figure 1. Bland- Altman plot of HR_{peak} GXT1 and HR_{peak} GXT2 in beats/min; n=26; LOA -16 to 14 (19.4% of mean); mean difference -1; Standard Deviation of the difference 8 beats/min

The LOA are determined in a Bland-Altman plot (Figure 1). The LOA for HR_{peak} of GXT1 and GXT2 range from -16 to14, which is 19% of the mean HR_{peak} (mean HR_{peak} = 155). This means that the individual variability represented in the region between the dotted lines has a width of 19% of the mean (Figure 1).

The achieved levels of GXT1 and GXT2 were normally distributed. The ICC between the attained levels was very good (ICC=0.95; 95%CI 0.90-0.98). The t-test showed no significant difference between the attained levels of GXT1 and GXT2 (p=0.56) (Table 3).

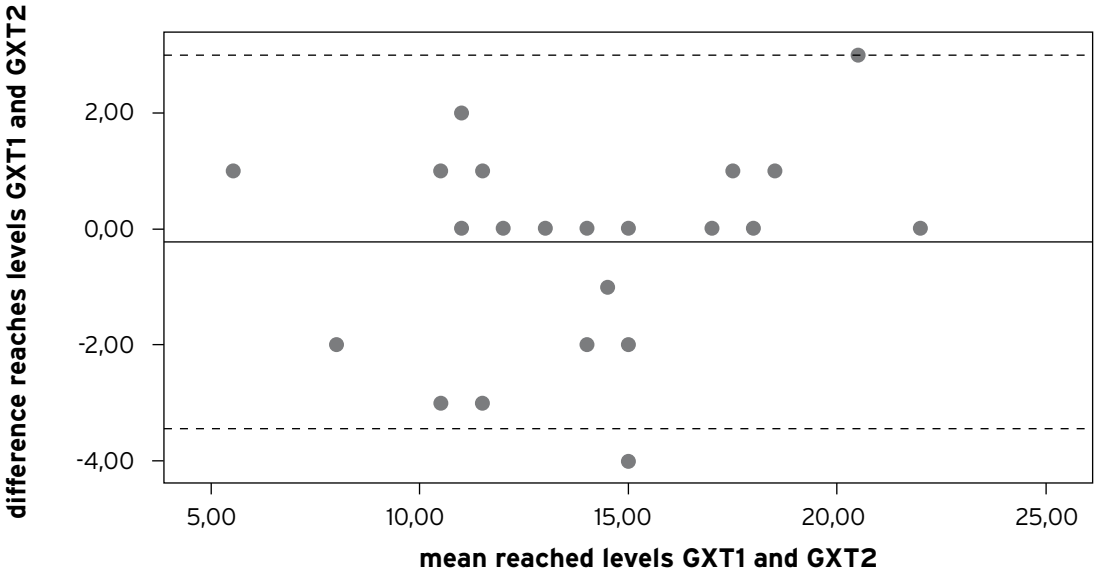


Figure 2. Bland and Altman plot of reached levels GXT1 and GXT2; n=26; LOA -3.4 to 2.9; (43.4% of mean); mean difference -0.19

The LOA for the attained levels of GXT1 and GXT2 range from -3.4 to 2.9 (Figure 2).
The mean level reached in both tests was 13.6.

The SDD was 2 levels indicating that persons had to increase or decrease their treadmill performance with 2 levels to have a relevant change in endurance time.
Table 3 presents the SEM, SDD and ES. The mean difference in HR_{peak} between GXT1 and GXT2 was 1 beat /minute. The mean difference in HR_{peak} between the SMBT and the corresponding GXT was also 1 beat /minute. These are both smaller than the SEM (5 beats/minute in both tests) (Table 3). The SDD for HR_{peak} was 6 beats/minute. Thus one can only speak of a significant difference when confronted with a difference larger than 6 beats/minute. The Effect Size (ES) for both HR_{peak} and the attained level ranges from small to moderate (0.12) between GXT1 and GXT2 [32].

Table 3. Test-retest reliability of HR_{peak} and levels achieved

	HR_{peak} GXT1 - HR_{peak} GXT2 (n=26)	HR_{peak} GXT - HR_{peak} SMBT (n=20)	HR_{peak} GXT1 - HR_{peak} calculated (n=26)	HR_{peak} GXT2 - HR_{peak} calculated (n=26)	HR_{peak} SMBT - HR_{peak} calculated (n=20)	HR_{peak} GXT1 - HR_{peak} GXT2 (n=26)
LOA **% of mean**	-16 to 14 19.5%	-17 to 15 20.5%	-45 to 13 35.8%	-43 to 13 34.5%	-42 to 14 34.1%	-3.4 to 2.9 43.4%
ICC **(95%CI)**	0.95 (0.88-0.98)	0.94 (0.86 – 0.98)	0.28 (-0.23 – 0.62)	0.39 (-0.21 – 0.72)	0.36 (-0.25- 0.72)	0.95 (0.90-0.98)
t-test	-0.6 (p=0.5)	-0.3 (p=0.7)	-5.4 (p=0.001)	-5.3 (p=0.001)	-4.6 (p=0.001)	-0.6 (p=0.56)
SEM	5	5	8	9	8	1
SDD	6	6	8	8	8	2
ES	0.12	0.07	0.73	0.73	0.73	0.12

HR: heart rate; GXT: graded exercise test; LOA: Limits of Agreement; ICC: Intraclass Correlation Coefficient; SEM: Standard Error of Measurement SEM=SDx√(1-ICC)[27]; SDD: Smallest Detectable Difference SDD=1,96x√(2xSEM) [27]; ES: Effect Size; ES=√{t^2/(t^2+df)} [32]

Validity

Twenty participants performed the SMBT, according to criterion A. The ICC (ICC=0.94; 95%CI 0.86-0.98) between the HR_{peak} SMBT and the corresponding HR_{peak} GXT was very good (Table 3). The t-test (p=0.7) revealed no significant differences between the HR_{peak} of the SMBT and the HR_{peak} of the corresponding GXT (Table 3). The LOA results range from -17 to 15. The LOA represents 20.5% of the mean (157 beats/min) (Figure 3).).

Eleven persons met the criterion of a maximal performance on the GXT, because the heart rate of the SMBT differed by not more than two beats from the HRpeak reached during the corresponding GXT. In four persons HRpeak SMBT was more than two beats more than HRpeak GXT, which means they did not meet the criterion of a maximal performance on the GXT. In five persons HRpeak SMBT was less than HRpeak GXT.

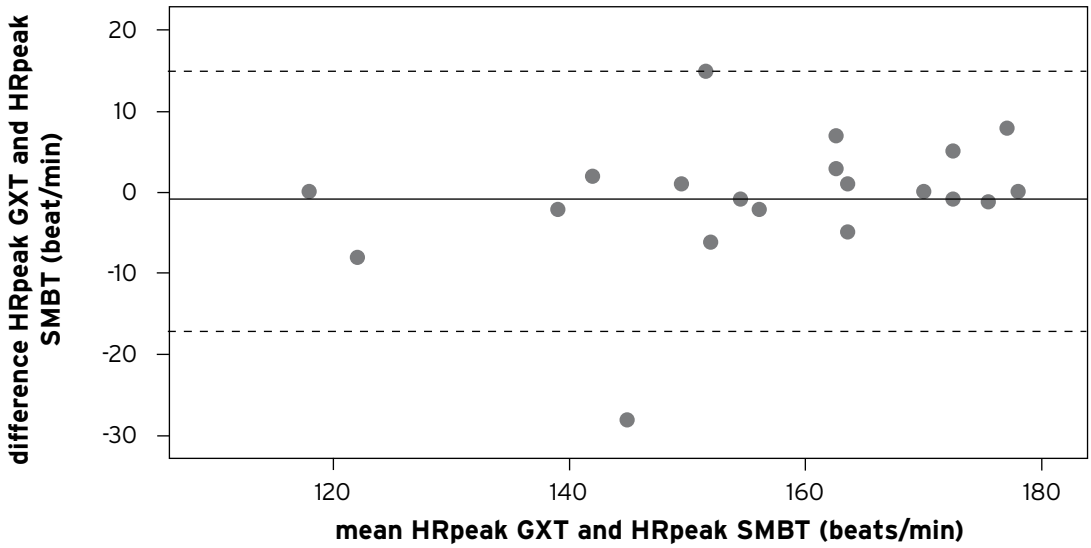


Figure 3. Bland-Altman plot of HR_{peak} GXT and HR_{peak} SMBT in beats/min; n=20; LOA -17 - 15; (20.5% of mean); mean difference -1; Standard Deviation of the difference 8 beats/min

However, the ICCs between the estimated HR_{peak} and HR_{peak} GXT1, HR_{peak} GXT2 and HR_{peak} SMBT were all poor (Table 3). The t-statistics comparing estimated HR_{peak} with HR_{peak} GXT1, HR_{peak} GXT2 and HR_{peak} SMBT all showed a significant difference (Table 3). This means that the agreement between estimated HR_{peak} against HR_{peak} GXT1, HR_{peak} GXT2, and HR_{peak} SMBT was poor (Table 3).

The ES for HR_{peak} between the SMBT and the corresponding GXT was small (0.07) [32]. The ES for HR_{peak} between GXT1, GXT2 and SMBT with the estimated HR_{peak} were all large (>0.5) [32]. This indicates a large difference between the measures for HR_{peak} found and estimated HR_{peak}', which corresponds to the significant difference between these measures recorded by the t-test.

Spearman's correlation of the VAS motivation scores between the instructor and test leader in GXT1 was low (rho 0.3; p=0.09) but significant for GXT2 (rho 0.7; p<0.001) and SMBT (rho 0.6; p=0.004). So, analysing a correlation between scored motivation and HR_{peak} would only be appropriate for GXT2 and SMBT. Spearman's correlation between scored motivation and HR_{peak} was significant only for GXT2 for both the instructor (p=0.02) and the test leader (p=0.03) (Table 4).

Table 4. Correlation scored motivation by Instructor or Test Leader and HR_{peak}

Correlation scored motivation by INS or TL and HR_{peak}						
	GXT1		**GXT2**		**SMBT**	
	score INS – HR_{peak}	score TL – HR_{peak}	score INS – HR_{peak}	score TL – HR_{peak}	score InS – HR_{peak}	score TL – HR_{peak}
Spearman rho	0.18 p=0.37 n=26	0.48 p=0.01 n=26	0.44 p=0.02 n=26	0.44 p=0.03 n=25	0.36 p=0.11 n=19	0.48 p=0.03 n=19
Critical value rho p=0.05; two-tailed	0.390	0.390	0.390	0.398	0.460	0.460

GXT: Graded Exercise Test; SMBT: Supra-Maximal Block Test; INS: Trained physical education instructor; TL: test leader; p: significance value; n: participants

All but one participant (96%) needed encouragement during GXT1 or GXT2. Fourteen (54%) needed encouragement during both tests and eleven participants (42%) needed encouragement during one test. Out of the total of fifty-two tests of GXT1 and GXT2, a lot of encouragement was given during twelve tests and some encouragement was given during twenty-seven tests.
Both GXT1 and GXT2 lasted 14 minutes (SD 3.8) on average, with a range of 4 to 22 minutes.

Correlation between HR_{peak} of GXT1 on the treadmill and HR_{peak} of the over ground test was poor and not significant (ICC=0.21; 95%CI -0.24–0.59). For the GXT2 the correlation was also poor and not significant (ICC=0.24; 95%CI -0.22–0.63).
The HR_{peak} for GXT1 and GXT2 treadmill were significantly higher than the tests performed over ground, as was expected.

Correlation between the levels of GXT1 treadmill and aSRT1 over ground was moderate but significant (ICC=0.67; 95%CI = 0.18-0.87). Correlation between the levels achieved for the GXT2 was moderate and also significant (ICC=0.73; 95% CI = 0.22-0.90). The wide confidence intervals indicate diversity among the participants. The achieved levels show a significant difference between the treadmill and the over ground test for the GXT1 (Wilcoxon signed ranks, p=0.03) and the GXT2 (Wilcoxon signed ranks, p=0.01).

Estimated peak heart rate

In our sample the ICC for HR_{peak} SMBT and the estimated HR_{peak} was poor (ICC=0.36; 95%CI -0.25–0.72). The t-statistic showed a significant difference (t=-4.6; p=0.001). Figure 4 shows a scatter plot of these findings. The equation seems to over-estimate HR_{peak} for people with SIMD. To adjust the equation to be relevant to our sample a new constant β=0.88 was calculated.

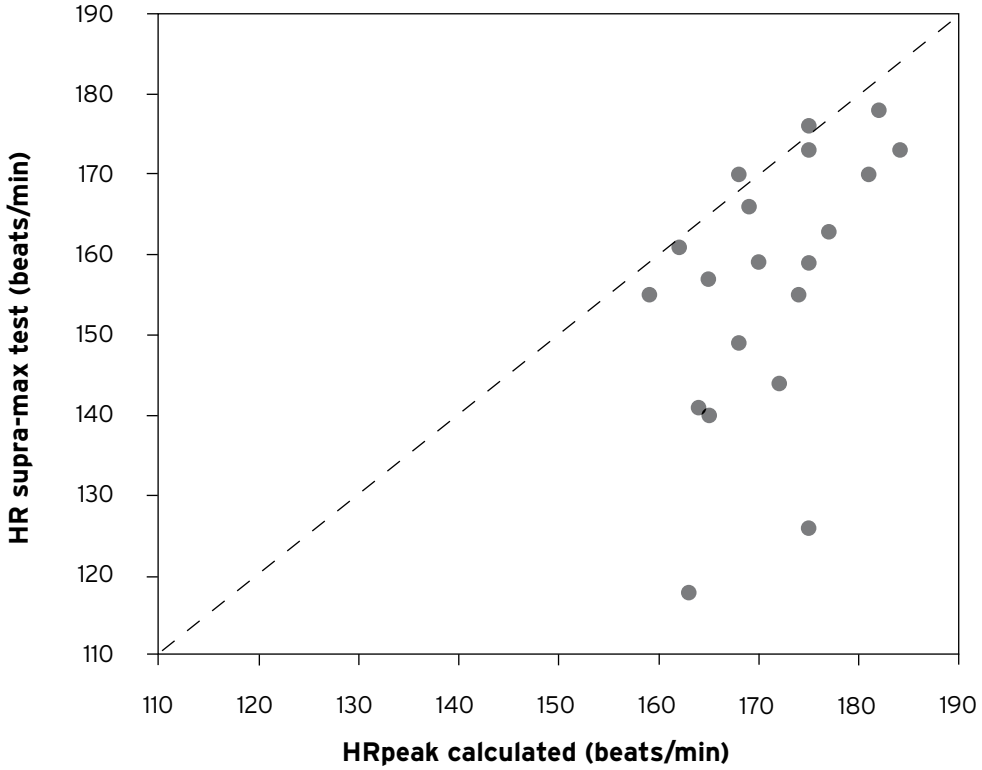


Figure 4 Scatter plot calculated HR_{peak} and HR_{peak} of supra-maximal block test (Pearson's r=0.5; p≤0.05); t-statistic p=0.0001). The dotted line representing perfect correlation, line of identity and absolute agreement.

Discussion

This study investigated the feasibility, test-retest reliability and validity of the adapted Shuttle Run Test protocol performed on a treadmill (graded exercise test, GXT) for people with SIMD and GMFCS-I. The results indicate that the GXT protocol performed on the treadmill has sufficient feasibility, based on the 80% criterion [28]. The test-retest reliability and validity of the GXT treadmill were sufficient for the sample. Furthermore, the results show that the Fernhall equation predicting HR_{peak} consistently overestimated the realized HR_{peak}.

High correlation (ICC=0.95; 95%CI 0.88–0.98) and good agreement (p=0.5) has been showed in this study between the GXT1 and GXT2 for HR_{peak}, indicating good test-retest reliability. This is in line with the results Fernhall et al. found in a reliability study performed with mentally retarded adults [33]. A SEM of 5 and an SDD of 6 with a small ES indicate that the margin for error is 5 and hence only a difference of more than 6 beats/minute indicates a significant difference. The results of the GXT1 and GXT2 are not significantly different and the ES are small. This confirms good test retest reliability.

A measure for variability is the LOA. With a 19.4% of the mean HR_{peak}, the LOA can be said to be wide. A wide LOA points at a considerable variation at the individual level and may be due to difficulties in motivation, behaviour or an inability to cope with stress, all of which are typically prevalent in people with SIMD [1, 3, 33].

In order to check the validity of the test protocol, the HR_{peak} SMBT was compared with the

preceding GXT results on HR_{peak}. Good correlation (ICC= 0.94; 95%CI 0.86-0.98) and agreement (p=0.7) for HR_{peak} exists between SMBT and the preceding GXT. These results might serve as indicative for a gold standard according to Midgley et al. [18, 19]. Midgley et al. reviewed available literature regarding the verification phase following a graded exercise test. In the present study, speed during the verification phase was 0.2 or 0.3 km/h higher than speed during the last completed stage, as in persons with SIMD only an increase of speed of 0.2 or 0.3 km/h is feasible [12]. This is considered an acceptable increase of speed, because Midgley et al [19] stated that 'the verification phase should incorporate a workload higher than that attained in the incremental phase to conform to the original concept of VO2max'. Furthermore, they indicated that the mean difference in HR_{peak} between the test and the SMBT should be maximal 1.9 beats/minute (SD 1.7) [19]. The results of our study show a mean difference of 1 beat/minute and a standard deviation of 1, which is within the limits drawn by Midgley [19]. The attained performance levels thus were reached using maximal effort. When looked at the individual level, eleven persons met the criterion of a maximal performance on the GXT, whereas four persons did not meet this criterion.

The actual and estimated HR_{peak} show low correlation. This could indicate that the equation developed by Fernhall et al. [13] is not applicable to our sample of participants. In the Fernhall study, 144 of the 276 participants had mild ID [13]. Similar to the findings of Kittredge et al. [34] we found the HR_{peak} to be significantly lower than the estimated HR_{peak} [34], suggesting that the formula's constant (β=0.56 in the Fernhall equation) should be corrected and take a higher value for people with SIMD. Hence, a recommendation for future research is to adjust the equation for estimating HR_{peak} for individuals with SIMD, enabling a valid calculation of the estimated HR_{peak}, which is crucial in assessing whether target heart rate has been reached for this specific group of particpants.

Since both HR_{peak} and attained levels are significantly higher in the treadmill test than in the over ground test, the treadmill test is valid for measuring maximal exercise capacity in people with SIMD and GMFCS-I. Environmental factors may explain the test results of the GXT protocol as performed on the treadmill [35]. The constraints-led approach limits the degrees of freedom during testing [35], resulting in a more restrictive test situation. The physical constraints of the treadmill, the sound of the running walking belt, the side bars and the instructor standing behind the participant all may stimulate the participant to continue walking.

Furthermore, issues of motivation, stress and the ability to understand test directions should be considered when interpreting test results [36]. For people with SIMD motivation for physical activity is low [36] and extrinsic encouragement and rewards often dictate activity performance [37]. This pattern of behavior is also evident in our study. All but one participant needed encouragement, fourteen during both test sessions and eleven during one of the test sessions.

Moreover, unfamiliar situations caused stress in several participants. Rintala et al. reviewed the familiarization process in cardiorespiratory fitness testing in persons with mild to moderate intellectual disabilities, recommending a familiarization protocol [21]. In our study the protocol consisted of walking the treadmill at least twice at regular walking speed before testing. Participants did not reach their HR_{peak} levels during these practice sessions. Frey et al. described how people with ID are hardly challenged by their support systems to exert physically [3]. During data collection most participants were challenged up to volitional exhaustion [16] for the first time in their lives. This may put forward an explanation for the insufficient feasibility of the SMBT. By letting future participants practice at a sub-maximal exercise level, the feasibility of the SMBT may improve. Furthermore, in future studies a familiarization protocol should be established with

well defined criteria for advancement from one familiarization level to the next.

Since people with intellectual disabilities tend to have lower motivation for physical activity [3, 16], we included a motivation score into the testprotocol. An aspect that may have influenced the motivation scores was that both the test leader and the instructor were aware of the estimated HR_{peak}. This may have influenced the encouragement given, and as a consequence, the scored motivation. A significant correlation between observed motivation and HR_{peak} existed during GXT2 for both the instructor and the test leader (Table 4). Nonetheless, in future studies the inter-observer reliability of scored motivation should be assessed as well.

Handrail support during steady-state treadmill exercise reduces the momentary aerobic demands [38, 39]. All but one participant held on to the handrail during the test procedure. When walking speed increased some participants leaned more heavily on the handrail which may have had an influence on the levels or HR_{peak} [39, 40]. The achieved levels in the GXT treadmill may have been relatively high as a consequence of leaning on the handrail.

A limitation of this study was how to decide when maximal exercise level or volitional exhaustion was reached. As for now, realizing the estimated HR_{peak} seems to be the only objective measure of maximal performance, which in the present study none of the participants achieved. Rintala et al. described volitional exhaustion [16] by signals such as heavy breathing, maximal heart rate, uncoordinated walking, sweating or verbal protest, which is too wide a range for a clear and workable measure. In future studies the volitional exhaustion has to be defined in a more accurate way.

The results of this project could be used to develop an experimental study investigating the trainability of exercise capacity in people with SIMD. Treadmill training could possibly improve health related physical fitness and thereby health for people with multiple disabilities. Further experimental research on training a population with SIMD is recommended.

Conclusion

The main conclusion of our study is that a GXT protocol performed on a treadmill is a feasible, reliable and valid test for determining HR_{peak} and exercise levels for people with SIMD and GMFCS-I. For this population, the GXT protocol has better validity for determining HR_{peak} and maximal level than the SRT over ground.

For future research, we recommend a revision of Fernhall's equation so as to enable a better prediction of the HR_{peak} for people with SIMD.

Furthermore, future studies should comprise of a familiarization protocol with well-defined criteria so as to reduce the influence of stress, stemming from unfamiliarity with the test situation, on the test results. Moreover, volitional exhaustion should be defined more clearly using unambiguous variables. Finally, an evaluation of inter-tester reliability of scored motivation should be established.

Acknowledgements

The authors kindly acknowledge and thank the participants for their participation in this study, their representatives for giving permission for this and the physical education instructors for their effort. The authors are grateful for the supportive and valuable feedback. The authors have no conflict of interest to declare.

References

1 Schalock R, Brown I, Brown R, Cummins RA, Felce D, Matikka L, Keith KD, Parmenter T. Conceptualization, measurement, and application of quality of life for persons with intellectual disabilities: report of an international panel of experts. Ment Retard. 2002;40(6):457-70.

2 Van den Broek EG, Janssen CG, van RT, Deen L. Visual impairments in people with severe and profound multiple disabilities: an inventory of visual functioning. J Intellect Disabil Res. 2006;50(Pt 6):470-5.

3 Frey GC, Buchanan AM, Rosser Sandt DD. "I'd rather watch TV": an examination of physical activity in adults with mental retardation. Ment Retard. 2005;43(4):241-54.

4 Graham A, Reid G. Physical fitness of adults with an intellectual disability: a 13-year follow-up study. Res Q Exerc Sport. 2000;71(2):152-61.

5 Houwen S, Visscher C, Lemmink KA, Hartman E. Motor skill performance of school-age children with visual impairments. Dev Med Child Neurol. 2008;50(2):139-45.

6 Hopkins WG, Gaeta H, Thomas AC, Hill PM. Physical fitness of blind and sighted children. Eur J Appl Physiol Occup Physiol. 1987;56(1):69-73.

7 Evenhuis HM, Sjoukes L, Koot HM, Kooijman AC. Does visual impairment lead to additional disability in adults with intellectual disabilities? J Intellect Disabil Res. 2009;53(1):19-28.

8 Bouchard C, Shephard RJ, Stephens T. Physical activity, fitness, and health : international proceedings and consensus statement. Human Kinetics Proceedings of the 2nd International Conference on Physical Activity, Fitness and Health. 1994, Champaign, Il.

9 Angermayr L, Melchart D, Linde K. Multifactorial lifestyle interventions in the primary and secondary prevention of cardiovascular disease and type 2 diabetes mellitus--a systematic review of randomized controlled trials. Ann Behav Med. 2010;40(1):49-64.

10 Thomas GN, Jiang CQ, Taheri S et al. A Systematic Review of Lifestyle Modification and Glucose Intolerance in the Prevention of Type 2 Diabetes. Curr Diabetes Rev. 2010;6(6):378-87.

11 Wallman K, Goodman C, Morton A, Grove R, Dawson B. Test-retest reliability of the aerobic power index component of the tri-level fitness profile in a sedentary population. J Sci Med Sport. 2003;6(4):443-54.

12 Waninge A, Evenhuis IJ, van Wijck R, van der Schans CP. Feasibility and reliability of two different walking tests in persons with severe intellectual and sensory disabilities. accepted 2010.

13 Fernhall B, McCubbin J A, Pitetti KH, Rintala P, Rimmer JH, Millar AL, De Silva A. Prediction of maximal heart rate in individuals with mental retardation. Med Sci Sports Exerc. 2001;33:1655-1660.

14 Froelicher VF, Jr, Thompson AJ Jr, Davis G, Stewart AJ, Triebwasser JH. Prediction of maximal oxygen consumption. Comparison of the Bruce and Balke treadmill protocols. Chest. 1975;68(3):331-6.

15 Millar AL, Fernhall B, Burkett LN. Effects of aerobic training in adolescents with Down syndrome. Med Sci Sports Exerc. 1993;25(2):270-4.

16 Rintala P, Dunn JM, McCubbin JA, Quinn C. Validity of a cardiorespiratory fitness test for men with mental retardation. Med Sci Sports Exerc. 1992;24(8):941-5.

17 Davids K, Button C, Bennett S. Dynamics of Skill Acquisition, a constraints-led approach. 2008, Human Kinetics, Champaign, IL.
18 Bouter LM, van Dongen MCJM, Zielhuis GA. Epidemiologisch onderzoek. vijfde, herziene druk ed. 2005, Houten: Bohn Stafleu van Loghum.
19 Midgley AW & Carroll S. Emergence of the verification phase procedure for confirming 'true' VO2max. Scand J Med Sci Sports. 2009;(19):313-22.
20 Dutch Society for Doctors in the Care for people with an Intellectual Disability (NVAZ) Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans, behavioural code for doctors in the assessment of resistance among people with an intellectual disability, 1999.
21 Rintala P, McCubbin JA, Dunn JM. Familiarization process in cardiorespiratory fitness testing for persons with mental retardation. Sports Med Train Rehabil. 1995;6(1):15-27.
22 Vashdi E, Hutzler Y, Roth D. Compliance of children with moderate to severe intellectual disability to treadmill walking: a pilot study. J Intellect Disabil Res. 2008;52(Pt 5):371-9.
23 Mc Cormack HM, Horne de DJL, Sheather S. Clinical applications of visual analogue scales: a critical review. Psych Med. 1988;18:1007-1019
24 Wewers ME, & Lowe NK. A critical Review of Visual Analogue Scales in the Measurement of Clinical Phenomena. Research in Nursing & Health. 1990;13:227-236.
25 Climstein M, Pitetti KH, Barrett PJ, Campbell KD. The accuracy of predicting treadmill VO2max for adults with mental retardation, with and without Down's syndrome, using ACSM gender- and activity-specific regression equations. J Intellect Disabil Res. 1993;37(Pt 6):521-31.
26 Fernhall B, Pitetti KH, Rimmer JH et al. Cardiorespiratory capacity of individuals with mental retardation including Down syndrome. Med Sci Sports Exerc. 1996;28(3):366-71.
27 Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26(4):217-38.
28 Lemmink KAPM. De Groninger Fitheidstest voor Ouderen. Ontwikkeling van een meetinstrument. 1996, Dissertation, University of Groningen, Netherlands.
29 Portney LG, Watkins MP. Foundations of clinical research. 2nd ed. Upper Saddle River, 2000, Prentice-Hall Inc, New Jersey.
30 Bland JM & Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986;1:307-310.
31 de Groot JF, Takken T, Schoenmakers MA, Tummers L, Vanhees L, Helders PJ. Reproducibility of energy cost of locomotion in ambulatory children with Spina Bifida. Gait Posture 2010;31(2):159-63.
32 Field A. Discovering Statistics Using SPSS. SAGE Publications Ltd; 2005.
33 Fernhall B, Millar AL, Tymeson GT, Burkett LN. Maximal exercise testing of mentally retarded adolescents and adults: reliability study. Arch Phys Med Rehabil. 1990;71(13):1065-8.
34 Kittredge JM, Rimmer JH, Looney MA. Validation of the Rockport Fitness Walking Test for adults with mental retardation. Med Sci Sports Exerc. 1994;26(1):95-102.
35 Newell KM. Constraints on the development of coordination. In Motor Development in Children: Aspects of Coordination and Control . In: M.G.Wade & H.T.A.Whiting, editor. Dordrecht, Netherlands: Martinus Nijhoff; 1986:341-60.
36 Pitetti KH, Rimmer JH, Fernhall B. Physical fitness and adults with mental retardation. An overview of current research and future directions. Sports Med. 1993;16(1):23-56.

37 Stanish HI, Temple VA, Frey GC. Health-promoting physical activity of adults with mental retardation. Ment Retard Dev Disabil Res Rev. 2006;12(1):13-21.
38 Rintala P, McCubbin JA, Dunn JM. Familiarization process in cardiorespiratory fitness testing for persons with mental retardation. Sports Med Train Rehabil. 1995;6(1):15-27.
39 Christman SK, Fish AF, Bernhard L, Frid DJ, Smith BA, Mitchell L. Continuous handrail support, oxygen uptake, and heart rate in women during submaximal step treadmill exercise. Res Nurs Health. 2000;23(1):35-42.
40 McConnell TR, Foster C, Conlin NC, Thompson NN. Prediction of Functional Capacity During Treadmill Testing: Effect of Handrail Support. J Cardiopulmonary Rehabil. 1991;11:255-60.

Chapter 6

Feasibility and reliability of a modified Berg Balance Scale in persons with severe intellectual and visual disabilities.

A. Waninge
R. van Wijck
B. Steenbergen
C.P. van der Schans

Journal of Intellectual Disability Research 2011;55(3):292-301

Abstract

Background The purpose of this study was to determine the feasibility and reliability of the modified Berg Balance Scale (mBBS) in persons with severe intellectual and visual disabilities (severe intellectual and multiple disabilities, SIMD) assigned Gross Motor Function Classification System (GMFCS) grade I and II.

Method Thirty-nine participants with SIMD and GMFCS grade I and II performed the mBBS twice with 1-week interval. Feasibility was assessed by the percentage of successful measurements per task and of the total score. First, test-retest reliability was determined by intraclass correlation coefficients (ICC) for each task and for the total score of all tasks combined. Second, level of agreement between test-retest scores was assessed with the proportion of equal scores for each task. Finally, internal consistency of the distinct tasks was assessed by Cronbach's alpha.

Results The results indicated that 92% of the measurements by the mBBS for all selected tasks were successful, indicating that the mBBS is a feasible instrument for the tested target group. ICC for the test-retest of the total score was 0.95. The proportion of equal scores for test-retest of the tasks was .80 or more, except for tasks 9 and 10. Cronbach's alpha of distinct tasks was 0.84. Test-retest reliability of tasks 9 and 10 was not acceptable.

Conclusions Feasibility of all tasks and test-retest reliability of 10 out of 12 mBBS tasks is acceptable. The mBBS is a both feasible and reliable test for evaluating the functional balance of persons with SIMD and GMFCS grades I and II.

Introduction

Locomotor skills in people with intellectual disabilities are characterised by decreased accuracy, variation and active exploration when compared to locomotor skills of those without intellectual disabilities [1]. Adults with mild or moderate intellectual disabilities are often found to have sensory integration problems [2] and a sedentary lifestyle [3, 4, 5]. IQ level is reported to be the main indicator of overall performance on motor tests [6]. Furthermore, a study by Wuang et al. [6] indicated that verbal comprehension and processing speed indexes specifically were reliable predictors of gross and fine motor function. Shinkfield et al. [7] reported that persons with intellectual disabilities suffer from inadequacies in both perception and motor-reproduction. Moreover, data on force platforms and posturography have outlined the characteristic movements of those with intellectual disabilities [8].

Like individuals with intellectual disabilities, persons with visual impairments also display poor performance on locomotor skills [9] and have low levels of habitual activity [10]. Compared with normal children, children with impaired vision exhibit differences in motor control which are not directly related to poor vision [11]. Reimer et al. [11] found that "children with visual impairment seemed to have more difficulties with calibrating the sensory information and specifically, they made larger errors along the lateral direction, when the target was not visible". As a result, persons with visual impairments often display poor physical fitness compared with persons with normal eyesight [12, 13].

Consequently, individuals that have both intellectual and visual disabilities are particularly at risk concerning the potential development of deficits in both locomotor skills as in daily functioning [14]. The high prevalence of visual impairment and blindness among persons with severe or profound intellectual disabilities suggests this risk is serious [15]. For complex reasons, individuals with intellectual disabilities frequently fall [16]. Visual deficits are identified as a potential factor for falling [16]. Furthermore, people with visual disabilities exhibit decreased balance [12, 13]. The combination of these findings puts forward the suggestion that persons having both intellectual and visual disabilities are likely to have decreased balance. It is imperative to gain insight into the severity and prevalence of balance problems in this population. In addition, interventions need to be designed to improve balance control, physical activity, and eventually, participation in daily life.

As to date, it remains unclear which specific balance test is feasible and reliable for testing subjects with severe intellectual and visual disabilities. It is certain, however, that many of the standardized outcome measures to quantify balance capabilities commonly used in physiotherapy are not applicable to participants with intellectual disabilities [16, 17]. If balance tests are to be used to assess persons with severe or profound intellectual and visual disabilities (severe multiple disabilities, SIMD), it follows that assessing the feasibility of these tests is a priority. If a participant does not understand the tasks of a certain test, the test will automatically fail to provide a realistic impression of the functional balance of the participant. In that case, the test will be invalid.

Therefore, test instructions for individuals with both intellectual and visual disabilities require our special focus. Two hindrances have to be taken into account. Firstly, as a result of severe or profound intellectual disability, test instructions are often not understood or with great difficulty (ICD-10, WHO) [16, 18]. Secondly, individuals with visual disabilities cannot see how test tasks are to be performed [15], rendering showing them how to perform the task at hand useless.

Out of the several balance tests described in the literature only a couple are feasible for our target group. Tests were assessed on the basis of the difficulty of test instructions and the functionality with regard to the target group. The following tests seemed adequate at first sight: the Functional Reach test [19], the Timed Up and Go Test (TUG) [20], the Performance Oriented Mobility Assessment (POMA) [21], the FICSIT-4 (Frailty and Injuries: Cooperative Studies of Intervention Techniques) [22] and the Berg Balance Scale (BBS) [23].

The Functional Reach Test [19] measures the difference between a subject's arm length and his/her maximal forward reach, as the subject sits or stands in a stationary position. For young subjects without disabilities, this test has a test-retest reliability of 0.89 and an interrater agreement of 0.98. Furthermore, this test is strongly associated with measurements of centre-of-pressure excursion, having a correlation coefficient of 0.71. However, after a few practice sessions the conclusion was reached that the Functional Reach Test is not suitable for the target group as they have difficulty understanding how to perform this task.

Podsiadlo and Richardson [20] modified the Get Up and Go Test [24] by incorporating a timed component and coined it the Timed Up and Go Test or TUG. In this test, the subject is observed and timed while he/she rises, walks, turns around and sits down again. The TUG has a test-retest reliability of 0.99 and an interrater agreement of 0.99. Moreover, TUG times correlate moderately well with the Barthel Index at 0.78 and scores on the BBS at 0.81. However, the speed of movement is influenced by a subject's comprehension time and reaction time, two factors that are inherently affected by intellectual and physical disabilities [1, 16]. Therefore, the abilities of persons with SIMD are underestimated if time is used as an outcome measure.

The POMA [21] evaluates balance when a subject stands, stands up, sits down, sits, and walks. Smits-Engelsman et al. [25] concluded that the sensitivity of the POMA is less than the sensitivity of the BSS, making the latter the preferred test.

The FICSIT-4 [22] comprises of four tests of static balance. These tests evaluate the ability to maintain balance in parallel, semi-tandem, tandem, and one-legged stances, alternately with eyes open and eyes closed. Test-retest reliability was good (r = .66) as was validity, showing moderate to high correlations with physical function measures and three balance assessment systems. However, after a few practice sessions it became clear that the subjects failed to understand the semi-tandem and tandem components of the FICSIT.

The BBS [23, 26, 27] evaluates a subject's functional balance during daily situations—such as when the subject stands up, stands still, sits down, picks something up from the ground, and turns around—using ratio scales when possible. The BBS has a test-retest reliability of 0.98 and an interrater agreement of 0.98. The BBS correlates well with the Barthel Index at 0.98 and with TUG scores at 0.70. This test has been proven to be sufficient for assessing different target populations, such as the elderly [28] and stroke patients [29]. The BBS was considered to be suitable for participants with SIMD because it assesses a person's functional balance during daily situations, which can be scored independently by observing the participant's spontaneous movements throughout the day. This way of scoring eliminates the risk of a patient not understanding the task. Yet, a few practice sessions showed some tasks to be too difficult for the participants, which led to a slight adaptation of the protocol by excluding four and adding two items. We coined the adapted BBS the modified Berg Balance Scale (mBBS). With these adaptations, the mBBS seems to be feasible for assessing balance in our target population.

To sum up, out of the five potential tests the literature search put forward, solely the BBS seems suitable for our target population, albeit only in its modified version. Hence, the purpose of

this study was to evaluate the feasibility and reliability of the modified Berg Balance Scale (mBBS) in persons with intellectual and visual disabilities classified Gross Motor Function Classification System (GMFCS) grade I and II.

Methods

Participants

Participants were recruited from a residential care facility for the profound or severe intellectually and visually disabled in the Netherlands. Of the residents of this facility, 65% also suffers from motor disabilities. The participants were classified according to their motor skills using the Gross Motor Function Classification System (GMFCS) [30], a five-level system used to classify the motor abilities of the physically disabled. Participants with a "Level I" classification can generally walk without restrictions but tend to have limitations in more advanced motor skills. Participants with a "Level II" classification can walk with slight restrictions and do not spontaneously increase their speed during walking. The locomotor skills of those assigned GMFCS levels III to V are very limited and they were therefore excluded from performing the balance test.

Written consent was requested from the representatives of 92 candidates and obtained from 80. After informed consent was obtained, the subjects were screened based on an examination by both a special needs physician and a behavioral scholar. The screening itself excluded seven subjects. Another eight participants were excluded, as they did not live at the centre where the tests were to be performed. Eight other participants were excluded as they could not attend all five practice sessions. Another 18 participants were excluded as they exhibited one or more of the exclusion criteria at the time of measurement (Fig. 1). These exclusion criteria were: psychoses, depression or other severe psychological problems; somatic diseases, which were defined as chronic diseases and/or diseases that do not resolve in the short term (e.g., osteoarthritis, osteoporosis, pneumonia, etc); general illness or fever; taking antibiotics; worsening of asthma, epilepsy (recent insult or epileptic fits), fresh wound(s)/bruise(s) or other factors causing pain during movement; and finally stress due to the participant's behavior shortly prior to the date of measurement.

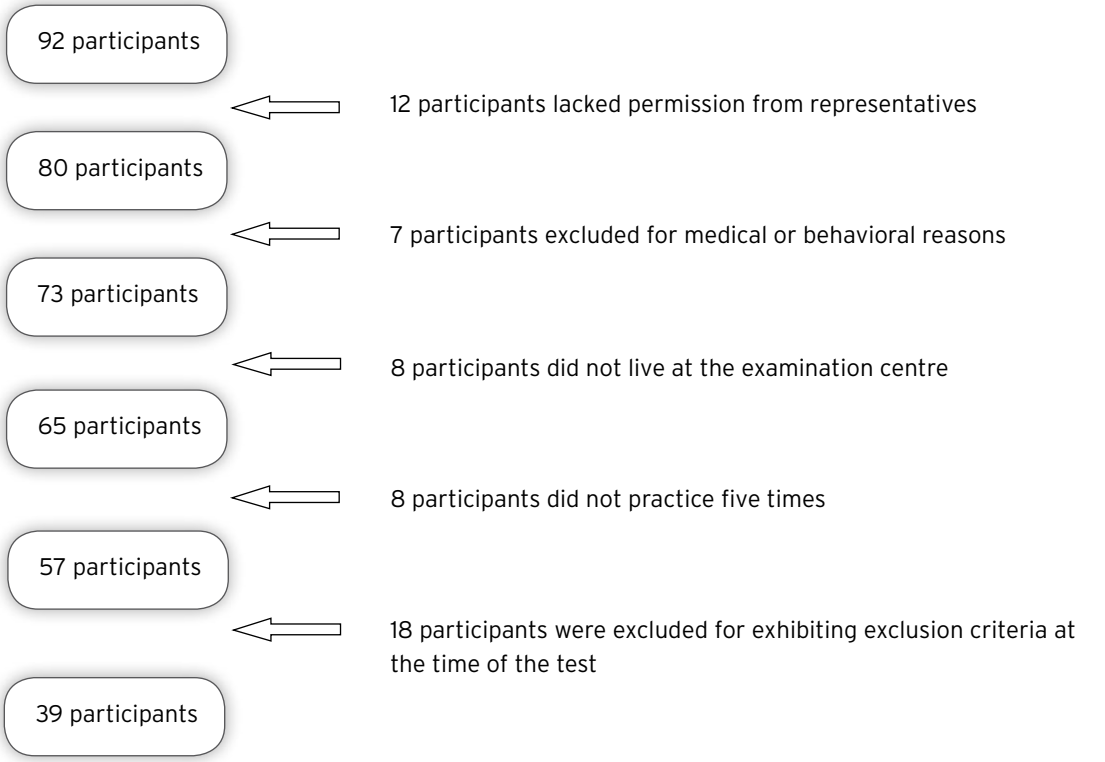


Figure 1. Inclusion steps.

Out of the remaining 39 participants in this study 28 were male and 11 were female. The mean (SD) age was 38 (11) years for the men and 44 (10) years for their female counterpart. Twenty-three participants were classified GMFCS level I and 16 participants GMFCS level II. According to the classification scheme of the ICD-10 (WHO, 1992), 92% (n=36) had severe intellectual disabilities, and 8% (n=3) suffered from profound intellectual disabilities. According to the WHO guidelines [31], all participants suffered from impaired vision: 44% (n=17) of the participants was severely partially sighted, 38% (n=15) was partially sighted and 18% (n=7) was slightly limited in sight. Most participants had impaired motor abilities: 67% (n=26) had orthopedic defects and 5% (n=2) had severe motor handicap of neurological origin. In addition, 23% (n=9) of the participants had slight hearing problems, 2% (n=1) had loss of hearing, and 5% (n=2) had either severe hearing loss or was completely deaf.

Design

The participants performed the mBBS twice, with a one-week interval between test and retest. Both tests were performed at the same time of day, under the same circumstances and under the supervision of the same personal caretaker and observer.

Ethical Statement

The study was performed in agreement with the guidelines of the Helsinki Declaration as revised in 1975. Permission was obtained from the institutional ethics committee. Informed consent was

obtained from the legal representatives of the participants, as the participants themselves were unable to give consent. The measurements were performed in accordance with the guidelines of the Dutch Society of Special Needs Specialists (NVAZ) which are outlined in a code called "Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans" [32]. This code intends to guide doctors in assessing resistance in persons with an intellectual disability. Following this code, consistent distress or unhappiness of the participant was interpreted as a sign of lack of assent, and further participation in the study was reconsidered.

Measures and Protocols

Prior to the measurements, the observers and personal caretakers of the participants completed a checklist that included all exclusion criteria. Participants were to be excluded from the study if they exhibited any of the exclusion criteria at the time of measurement.

As familiarisation, the participants practiced five times prior to formal testing. As Hale et al. [16] have noted, allowing a participant to become familiarized with both test and tester may ease problems concerning misunderstanding of the required tasks ahead. In these practice sessions the following tasks were found too difficult for the participants to perform: tandem standing, reaching forward while standing, turning one's trunk while feet are fixed and standing with eyes closed. Therefore, the protocol was slightly adapted by excluding these four components while adding two new items: walking on a thin line and walking on a gymnastic beam (width 30 cm, 40 cm above the floor). These two items were added since the participants were already familiar with these tasks. We coined the adapted BBS the modified Berg Balance Scale (mBBS). Including the aforementioned adaptations, the feasibility and test-retest reliability of the mBBS was examined. The mBBS consisted of 12 items, as shown in Table 1. The performance on each of these items was scored on a 5-point ordinal scale (0-4 points), where a score of 0 denotes the inability of the participant to perform the task, and a score of 4 is assigned when the participant is able to complete the task based on the criterion that has been assigned to it. The maximum score of the mBBS is 48 points. If a subject did not understand a task, the score of that task was excluded from the total score.

During testing, two observers completed the score forms independently and a personal caretaker instructed the participants. In total, two observers and four caretakers participated in the study. The observers were physiotherapist students, who performed the study for their bachelor thesis and were supervised by the first and second author. All observers and caretakers were instructed during two separate training sessions so as to ensure consistency among them. The first training session was supervised by the first and second author and took 2 h. The protocol of the original BBS was the topic of the first training session and a detailed manual was provided to each observer. During the five aforementioned practice sessions, both the observers and caretakers practiced using the instructions and scoring procedures. The scoring procedure was accurately determined and the scores of the two observers were compared. The level of consistency appeared to be sufficient. After the aforementioned adaptations of the BBS protocol, the second training session was organised with the adapted protocol, which was supervised by the first and second author too. This training session focused on the two new test items.

Table 1. The 12 items of the mBBS*

Number	Test item
1.	Sitting unsupported
2.	Change of position: sitting to standing
3.	Change of position: standing to sitting
4.	Transfers
5.	Standing unsupported
6.	Standing with feet together
7.	Turning 360 degrees
8.	Retrieving objects from floor
9.	Stool stepping
10.	Walking on a thin line
11.	Standing on one leg
12.	Walking on a gymnastic beam

mBBS, modified Berg Balance Scale

Data analyses

The data were analyzed using SPSS 14.0.

Feasibility

To assess feasibility, we held the number of successful measurements per task against the total number of measurements. As it only makes sense to use a test if there is a reasonable percentage of successful measurements, feasibility was considered to be sufficient if 85% of the measurements were successful [33].

Test-retest reliability

To determine the test-retest reliability, we computed intraclass correlation coefficients (ICC; two-way random, absolute agreement). Reliability was considered to be moderate, if the ICC was between .41 and .60, strong if the ICC was between .61 and .80, good if the ICC was greater than .81 [34] and very good if the ICC was greater than .90 [35]. To assess the internal consistency between the 12 test tasks, we computed Cronbach's alpha. Internal consistency was acceptable if Cronbach's alpha was .70 or more [36]. To analyse the level of agreement between the scores for test and retest of the distinct tasks, the proportion of equal scores and its Wilson confidence intervals (CI) were computed, as suggested by Brown et al [37]. The level of agreement was considered to be sufficient if the proportion of equal scores is .80 or more and the Wilson CI (95 % CI) are between .60 and 1.0. Furthermore, we computed the power of the study with the hypothesis that the population proportion of agreement is 0.50, taking a one-sided test and sample size 39 [38, 39].

Modified Berg Balance Scale scores

In the BBS, a score of 80% (45 points) indicates sufficient balance [26]. However, this cut-off value cannot simply be applied to the mBBS, as the modifications influence the cut-off value and thus render a comparison meaningless. For that reason, we will describe the scores of the participants without the cut-off values.

Results

Feasibility

Tasks 1, 3, 4, 7, and 8 were completed by all 39 participants; tasks 2, 5 and 12 by 38 participants, tasks 9 and 11 by 37 participants, and tasks 6 and 10 by 36 participants (Table 2). Thirty-six out of 39 participants (=92%) completed all tasks. The duration of the test was about 30 minutes.

Table 2. Percentage successful mBBS measurements in GMFCS level I and II participants (n=39)*

	Tasks 1, 3, 4, 7, 8	**Tasks 2, 5, 12**	**Tasks 9, 11**	**Tasks 6, 10**
Percentage successful measurements	100%	97%	95%	92%

*mBBS, modified Berg Balance Scale; GMFCS, Gross Motor Function Classification System

Test-retest reliability

Table 3 summarizes the medians of test and retest, the results of the ICC analysis, and the proportion of equal scores.

The ICC for the tasks 1, 4, and 5 was considered moderate, for the tasks 2, 7, 9, 10 strong, for the tasks 6, 8, 11, 12, and the total score very good, whereas the ICC for the task 3 could not be computed, because the scale has zero variance items. The ICC for the total score without tasks 9 and 10 was 0.97 (0.94-0.98), which is very good [35]. Cronbach's alpha for tasks 1–12 was 0.84 [36]. The obtained proportions of equal scores were greater than or equal to 0.80 with Wilson 95% CI between 0.60 and 1.00 for tasks 1, 2, 3, 4, 5, 6, 7, 8, 11, and 12 (Table 3). However, the proportion of equal scores was <.80 for tasks 9 and 10 and the Wilson CI were wider than 0.40 for these tasks too.

The power analysis revealed a power of .91 with the hypothesis that the population proportion of agreement is 0.50 and the alternative and true agreement is 0.75 [38, 39](Brown et al, 2002; Dorai-Raj 2009).

Table 3. Medians of test and retest, results of the ICC analysis with 95% confidence intervals, and the proportion of equal scores with Wilson confidence intervals.*

	Median Test	Median Retest	ICC 95% CI	P of equal scores 95% CI
Task 1	4	4	0.68 0.40-0.84	0.82 0.67-0.91
Task 2	4	4	0.76 0.53-0.88	0.84 0.69-0.92
Task3	4	4	Could not be computed	0.97 0.87-1.00
Task 4	4	4	0.45 -0.46-0.71	0.90 0.76-0.96
Task 5	4	4	0.64 0.31-0.81	0.84 0.70-0.93
Task 6	0	0	0.91 0.81-0.95	0.88 0.72-0.95
Task 7	4	4	0.74 0.49-0.86	0.90 0.76-0.96
Task 8	2	2	0.99 0.98-0.99	0.92 0.79-0.97
Task 9	3	3	0.86 0.72-0.93	0.68 0.51-0.81
Task 10	2	2	0.72 0.48-0.86	0.54 0.39-0.68
Task 11	0	0	0.95 0.89-0.97	0.83 0.67-0.92
Task 12	4	4	0.98 0.95-0.99	0.94 0.82-0.98
Total score	35	35	0.95 0.92-0.98	-

*Two-way random; total agreement; ICC, Intra Class Correlation Coefficient, CI, confidence intervals; P, proportion.

Modified Berg Balance Scale scores

The median score of the GMFCS level I participants was 36 (24-47) and that of the GMFCS level II participants 29 (17-44).

Discussion

As to date, it is unclear which specific balance test can be feasibly and reliably used for individuals with SIMD. The results of the present study show that the feasibility of all mBBS tasks was acceptable for participants with SIMD and GMFCS level I and II. The test-retest reliability assessed with the ICC, was acceptable for 7 of the 12 mBBS tasks (indicating strong to very good reliability), as was the total score. Of the tasks 1, 4, and 5 the ICC was moderate. The level of agreement assessed with proportion of equal scores was acceptable (higher than .80) for 10 out of 12 mBBS tasks. The proportion of equal scores for tasks 9 and 10 was lower than .80. Taken the ICC and the proportion of equal scores together, we consider the reliability of the mBBS sufficient, except for tasks 9 and 10. The reliability of the total score computed with ICC when correcting for tasks 9 and 10 was very strong. Internal consistency between the tasks was acceptable.

The mBBS appears to be a feasible and suitable test, given the challenges in obtaining test results from participants with severe intellectual and visual disabilities [16].

The reliability of 10 of the 12 mBBS tasks (ICC 0.97) was comparable to the reliability of corresponding BBS tasks reported in other studies with other populations: a very good intrarater reliability was found by Berg et al. [26] in the elderly (ICC 0.97) and by Listen and Brouwer [40] on stroke patients too (ICC 0.98). The study of Blum & Korner-Bitensky [29] on stroke patients, reported an ICC of 0.97 for test-retest reliability. This is considered a satisfactory result for test-retest reliability too, given the aforementioned difficulties in obtaining test results from participants with SIMD [16].

In the present study, Cronbach's alpha was 0.84 for the mBBS, which is less reliable than the Cronbach's alpha of 0.98 reported by Blum & Korner-Bitensky [29]. Nonetheless, our Cronbach's alpha value is still within the acceptable range, according to Field [36].

The proportion of equal scores for the test-retest of task 9 was relatively low, which might be explained by the fact that the subjects had trouble understanding task 9, which involved stepping on to a seat. It was observed that subjects, placing one foot on the seat, either intuitively placed the other foot on the seat next to their first foot as if climbing stairs, or intuitively stepped over the seat. Task 11, standing on one leg, which is also included in FICSIT-4 [22], could act as a feasible and reliable alternative for task 9, as both tasks require a subject to stand on one leg.

Also the proportion of equal scores for the test-retest of task 10 was relatively low. Performing the task of walking on a thin line proved difficult for the subjects, who often were not able to see the line. It was tried to solve this problem by replacing the line by a thin rope, but it was found that the subjects still did not manage to complete the task. Task 12, walking on a gymnastic beam (width 30 cm, 40 cm above the floor), could act as a feasible and reliable alternative. Participants are more familiar with this task and it would therefore ease problems concerning understanding. Considering these observations, we recommend excluding tasks 9 and 10 because of their low proportion of equal scores and their relatively low percentage of successful measurements. The ICC of tasks 1, 4, and 5 was moderate, although the proportion of equal scores was acceptable. Furthermore, these tasks also proved to be feasible. Taking these findings into consideration, we recommend sustaining tasks 1, 4 and 5 in the mBBS. Consequently, the final mBBS consists of 10 tasks.

According to Berg et al. [41], the BSS cannot reliably estimate the probability of falling. For that reason, we propose to use the mBBS for evaluating the effects of intervention on balance. However, for this purpose, future research should aim to examine the sensitivity to change of the mBBS. We have the impression that the mBBS has floor and ceiling effects, implying that the mBBS may not always detect meaningful changes when evaluating an intervention. These effects are also described by Blum & Korner-Bitensky [29]. However, there were differences between the median scores of the GMFCS level I and II participants, 36 to 29, respectively. This might be indicative of the potentials of the mBBS to be discriminative. Further research on this topics may be useful.

A rather small number of participants participated in the present study, which could be a limitation. However, given the width of the Wilson confidence intervals of the proportion of equal scores, the power is sufficient except for Task 9 and 10 [38, 39]. Furthermore, our power analysis revealed a sufficient power of 0.91, with the hypothesis that the population proportion of agreement is 0.50, the alternative and true hypothesis is 0.75, taking a one-sided test and sample size 39 [38, 39]. Although 65 subjects were initially included in the study, only 39 met all inclusion criteria and were able to complete both test and retest. Some were excluded because they were unable to perform the mBBS test and retest within one week, others because they exhibited one or more of the exclusion criteria during retesting.

In conclusion, the results show that the mBBS is both a feasible and reliable test for evaluating the functional balance of individuals with severe intellectual and visual disabilities. Even though the Berg Balance Scale is widely used, its reliability for individuals with SIMD had not yet been evaluated. This research extends the knowledge for researchers and clinicians in the field using the BBS. As mentioned, using a modified version of the BSS rendered the standard BBS' cut-off scores meaningless. Further research should aim to develop cut-off values for the mBBS, to examine the validity of the mBBS, including the sensitivity to change, and the presence of floor and ceiling effects. Furthermore, research focused on the development of interventions aimed at improving balance control in persons with SIMD is recommended.

Acknowledgements

The authors kindly acknowledge and thank the participants for their participation in this study, their representatives for giving permission for this and the gymnastic instructors of Royal Dutch Visio The Brink for practising with the participants during the practice sessions and for accompanying the participants during the tests. Furthermore, the authors kindly acknowledge W. Krijnen, PhD, for helping with statistical analyses.

References

1 Van Erkelens-Zwets JHJ & Kars H. Problemen met het gedrag van zwakzinnige mensen. 1988, Samsom Stafleu, Alphen aan den Rijn.

2 Carmeli E, Bar-Yossef T, Ariav C, Paz R, Sabbaq H, Levy R. Sensorimotor impairments and strategies in adults with intellectual disabilities. Motor Control. 2008;12(4):348-61.

3 Frey GC & Chow B. Relationship between BMI, physical fitness, and motor skills in youth with mild intellectual disabilities. Int J Obes. 2006;30:861-867.

4 Lahtinen U, Rintala P, Malin A. Physical performance of individuals with intellectual disability: a 30 year follow up. Adapt Phys Act Q. 2007;24: 125-143.

5 Temple VA, Frey GC, Stanish HI. Physical activity of adults with mental retardation: review and research needs. Am J Health Prom. 2006;21:2-12.

6 Wuang YP, Wang CC, Huang MH, Su CY. Profiles and cognitive predictors of motor functions among early school-age children with mild intellectual disabilities. J Intellect Disabil Res. 2008;52(12):1048-60.

7 Shinkfield AJ, Sparrow WA, Day RH. Visual discrimination and motor reproduction of movement by individuals with mental retardation. American Journal of Mental Retardation. 1997;102(2):172-81.

8 Bodfish JW, Parker DE, Lewis MH, Sprague RL, Newell KM. Stereotypy and motor control: differences in the postural stability dynamics of persons with stereotyped and dyskinetic movement disorders. Am J Ment Retard. 20011;06(2):123-34.

9 Houwen S, Visscher C, Lemmink KAPM, Hartman E. Motor skill performance of school-age children with visual impairments. Dev Med and Child Neur. 2008;50(2):139-45.

10 Hopkins WG, Gaeta H, Thomas AC, Hill PM. Physical fitness of blind and sighted children. Eur J of Appl Phys Occ Phys. 1987;56: 69-73.

11 Reimer AM, Cox RF, Boonstra NF, Smits-Engelsman BC. Effect of visual impairment on goal-directed aiming movements in children. Dev Med Child Neur. 2008;50(10):778-83.

12 Häkkinen A, Holopainen E, Kautiainen H, Sillanpää E, Häkkinen K. Neuromuscular function and balance of prepubertal and pubertal blind and sighted boys. Acta Paediatrica. 2006;95(10):1277-83.

13 Seemungal BM, Glasauwer S, Gresty MA, Bronstein AM. Vestibular perception and navigation in the congenitally blind. J of Neurophys. 2007;97(6):4341-56.

14 Evenhuis HM, Sjoukes L, Koot HM, Kooijman AC. Does visual impairment lead to additional disability in adults with intellectual disabilities? J Intellect Disabil Res. 2009;53(1):19-28.

15 Van Splunder J, Stilma JS, Bernsen RM, Evenhuis HM. Prevalence of visual impairment in adults with intellectual disabilities in the Netherlands: cross-sectional study. Eye (Lond) 2006;20(9):1004-10.

16 Hale L, Bray A, Littmann A. Assessing the balance capacities of people with profound intellectual disabilities who have experienced a fall. J Intellect Disabil Res. 2007;51(Pt 4):260-8.

17 Hilgenkamp TIM, Van Wijck R, Evenhuis HM. Physical fitness in older people with ID—Concept and measuring instruments: A review. Res Dev Dis. 2010;31:1027–1038.

18 World Health Organization. International statistical classification of diseases and related health problems: tenth revision (ICD-10). 1992;369–370.

19 Duncan PW, Weiner DK, Chandler J, Studenski S. Functional reach: a new clinical measure of balance. J Gerontol. 1990;45(6):192-197.
20 Podsiadlo D. & Richardson S. The timed 'up and go: a test of basic functional mobility for frail elderly persons. Journal of the American Geriatrics Society. 1991;39:142-148
21 Tinetti M. Performance oriented assessment of mobility problems in elderly patients. Journal of the American Geriatrics Society. 1986;34:119-126.
22 Rossiter-Fornoff JE, Wolf SL, Wolfson LI, Buchner DM. A cross-sectional validation study of the FICSIT common data base static balance measures. Frailty and Injuries: Cooperative Studies of Intervention Techniques Journals of Gerontology. Series A: Biological Sciences and Medical Sciences. 1995;50(6).
23 Berg K. Balance and its measure in the elderly: a review. Phys Can. 1989;41:240-6.
24 Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the "Get-up and Go" test. Arch Phys Med Rehab. 1986;67:387-389.
25 Smits-Engelsman BCM, Bekkering GE, Hendriks HJM. (Eds.). Fysiotherapie and osteoporose. Amersfoort / Houten. Koninklijk Nederlands Genootschap voor Fysiotherapie. 2000, Bohn Stafleu Van Loghum (KNGF richtlijnen).
26 Berg K, Wood-Dauphinee S, Williams JI, Maki B. Measuring Balance in the Elderly: Preliminary development of an Instrument. Physiotherapy Canada. 1989;41:304-311.
27 Berg K, Wood-Dauphinee S, Williams JI, Gayton D. Measuring Balance in the Elderly: Validation of an Instrument. Can J Public Health. 1992;Supp2:07-11.
28 Steffen TM, Hacker TA, Mollinger L. Age – and Gender-Related Test Performance in Community-Dwelling Eldery People: Six-Minute Walk Test, Berg Balance Scale, Timed Up And Go Test, and Gait Speed. Physical Therapy. 2002;82:128-136.
29 Blum L & Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. Physical Therapy. 2008;88(5):559-66.
30 Palisano R, Hanna SE, Rosenbaum PL, Rusell DJ, Walter SD, Wood EP, Raina PS, Galuppi BE. Validation of a model of Gross Motor Function for Children With Cerebral Palsy. Physical Therapy. 2000;80:974-985.
31 World Health Organization (WHO). International Classification of Functioning, Disability and Health. Geneva, 2001.
32 Dutch Society for Doctors in the Care for people with an Intellectual Disability (NVAZ) Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans, behavioural code for doctors in the assessment of resistance among people with an intellectual disability, 1999.
33 Malmberg JJ, Miilunpalo SI, Vuori IM, Pasanen ME, Oja P, Haapanen-Niemi NA. A health-related fitness and functional performance test battery for middle-aged and older adults: feasibility and health-related content validity. Arch Phys Med. 2002;83(5):666-677.
34 Feinstein AR. Clinimatrics. 1987, New Haven, Yale University Press, 184-185.
35 Portney LG, Watkins MP. Foundations of clinical research. 2nd ed. Upper Saddle River, 2000, Prentice-Hall Inc, New Jersey.
36 Field A. Discovering Statistics Using SPSS. SAGE Publications Ltd; 2005.
37 Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. Statist Sci. 2001;16:101-133.
38 Brown LD, Cai TT, DasGupta A. Confidence Intervals for a Binomial Proportion 377 and Asymptotic Expansions. Ann Stat. 2002;30:160-201.

39 Dorai-Raj S. Binomial Confidence Intervals For Several Parameterizations. R Package 387 version 1.0-5. http://CRAN.R-project.org/package=binom (retrieved May 26th, 2009).
40 Liston RA & Brouwer BJ. Reliability and validity of measures obtained from stroke patients using the Balance Master. Archives of Physical Medicine and Rehabilitation. 1996;77: 425-430.
41 Berg K, Chesworth B, Muir SW, Speechley M. Use of the Berg Balance Scale for predicting multiple falls in community-dwelling elderly people: a prospective study. Physical Therapy. 2008;88(4):449-59.

Chapter 7

Feasibility, test-retest reliability and interrater reliability of the Modified Ashworth Scale and Modified Tardieu Scale in persons with profound intellectual and multiple disabilities.

A. Waninge
R.A. Rook
A. Dijkhuizen
E. Gielen
C.P. van der Schans

Research in Developmental Disabilities 2011;32(2):613-620
Reproduced by courtesy of Elsevier

Abstract

Caregivers of persons with profound intellectual and multiple disabilities (PIMD) often describe the quality of the daily movements of these persons in terms of flexibility or stiffness. Objective outcome measures for flexibility and stiffness are muscle tone or level of spasticity. Two instruments used to grade muscle tone and spasticity are the Modified Ashworth Scale (MAS) and the Modified Tardieu Scale (MTS). To date, however, no research has been performed to determine the psychometric properties of the MAS and MTS in persons with PIMD. Therefore, the purpose of this study was to determine the feasibility, test-retest reliability, and interrater reliability of the MAS and MTS in persons with PIMD. We assessed 35 participants on the MAS and MTS twice, first for the test and second a week later for the retest. Two observers performed the measurements. Feasibility was assessed based on the percentage of successful measurements. Test-retest and interrater reliability were determined by using the Wilcoxon signed rank test, intraclass correlation coefficients (ICC), Spearman's correlation, and either limits of agreement (LOA) or quadratically weighted kappa. The feasibility of the measurements was good, because an acceptable percentage of successful measurements were performed. MAS measurements had substantial to almost perfect quadratically weighted kappa (>0.8) and an acceptable ICC (>0.8) for both test-retest and interrater reliability. However, MTS measurements had insufficient ICCs, Spearman's correlations, and LOAs for both test-retest and interrater reliability. Our data indicated that the feasibility of the MAS and MTS for measuring muscle tone in persons with PIMD was good. The MAS had sufficient test-retest and interrater reliability; however, the MTS had an insufficient test-retest and interrater reliability in persons with PIMD. Thus, the MAS may be a good method for evaluating the quality of daily movements in persons with PIMD. Providing test administrators with training and clear instructions will improve test reliability.

Introduction

Persons with profound intellectual and multiple disabilities (PIMD) generally have very limited mobility, always use a wheelchair [1], and often have a Gross Motor Function Classification System (GMFCS) level of IV or V [2]. Spasticity, dyskinesia, or ataxia with hypotony frequently occurs in individuals in these GMFCS levels [3]. Lance [4] defined spasticity as "a motor disorder, characterised by a velocity-dependent increase in tonic stretch reflexes (muscle tone) with exaggerated tendon jerks, resulting from hyper-excitability of the stretch reflex as one component of the upper motor neurone syndrome". Muscle tonus or muscle tone is "the state of activity or tension of a muscle beyond that related to its physical properties, that is, its active resistance to stretch. In skeletal muscle, tonus is dependent upon efferent innervation" [5].

Persons with PIMD are at risk for a variety of limitations in daily functioning [6], such as inactivity, unsteady movement, and diminished initiative. However, research into the quality of daily movements of persons with PIMD is limited, and related knowledge is scarce. Caregivers of persons with PIMD often describe the quality of daily movements in terms of flexibility or stiffness. Objective outcome measures for flexibility and stiffness are muscle tone or level of spasticity.

Bohannon and Smith [7] introduced the Modified Ashworth Scale (MAS) as a scale for grading spasticity. The MAS is a clinical measure of muscle tone and a nominal-level measure of resistance to passive movement [8]. The reliability of the scale appears to be better for measuring muscle tone of upper limbs [8]. Although one study found the reliability of the MAS to be very good (kappa was .84 for interrater and .83 for intrarater comparisons) [9], other studies found it to be insufficient [10, 11, 12, 13].

Haugh, Pandayan, and Johnson [14] suggested that the Modified Tardieu Scale (MTS) is a more appropriate clinical measure of spasticity than the MAS. The MTS assesses resistance to passive movement at both slow and fast speeds, and therefore adheres more closely to Lance's definition of spasticity [4, 14]. Both parameters of the MTS have excellent test-retest and interrater reliability in children with cerebral palsy [15]. However, as with the MAS, other studies found the MTS to have insufficient reliability [12, 16, 17].

Both the MTS and the MAS show sufficient test-retest and interrater reliability in adults with intellectual disabilities [18], but the MTS seems to be more feasible and reliable than the MAS. The MTS also shows more reliability than the MAS in adults with severe brain injury and spasticity [19. Haugh et al. [14] stated that further studies need to be undertaken to clarify the validity and reliability of the MTS and the MAS for a variety of muscle groups in adult neurological patients. Thus far, no research has been performed to determine the psychometric properties of the MTS and MAS in persons with PIMD. Therefore, the purpose of this study was to determine the feasibility, test-retest reliability, and interrater reliability of the MAS and MTS in persons with PIMD.

Methods

Participants

We asked the representatives of 42 persons with PIMD for written permission for these persons to participate in our study. Forty representatives gave permission. After informed consent was obtained, the subjects were screened based on an examination by both a special needs physician and a behavioral scholar. The screening exclusion criteria were severe psychological problems

or somatic diseases, which were defined as chronic diseases and/or diseases that do not resolve in the short term. Two persons were excluded because they exhibited one of these problems or diseases. The exclusion criteria at the time the measurements were being performed were general illness or fever; taking antibiotics; recently started taking muscle relaxants; worsening of asthma, epilepsy (recent insult or epileptic fits); fresh wound(s)/bruise(s) or other factors causing pain during movement; or stress due to the subject's behavior just before the measurement date. Three persons were excluded because they exhibited one of these criteria. Figure 1 presents the sampling scheme of persons included in the study.

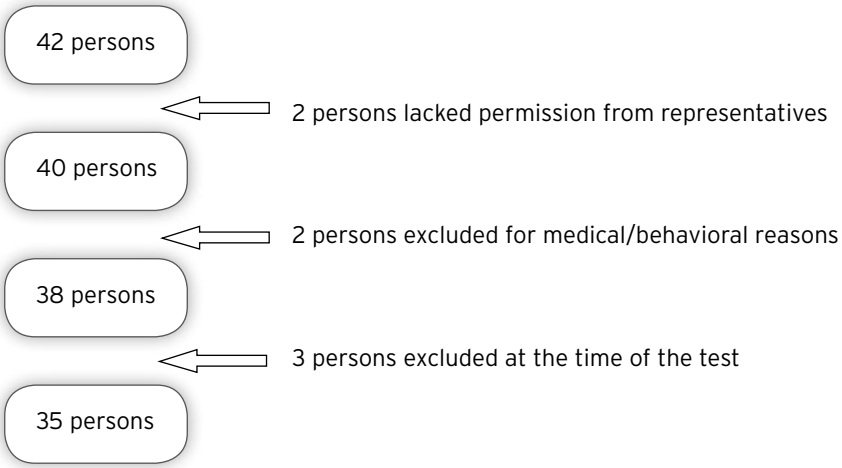


Figure 1. Sampling scheme of subjects included in the study.

The participants with PIMD were classified as GMFCS IV or V [2]. Furthermore, the intellectual level or intelligence quotient (IQ) of each participant was classified according to the International Classification of Diseases (ICD-10) of the World Health Organization (WHO) [20]. The presence or absence of epilepsy was also recorded, because we assumed that seizures greatly affect muscle tone. We also classified the visual impairments of the participants according to WHO guidelines [21]. Finally, the presence or absence of orthopedic disorders was recorded.

Ethical statement

The study was performed in agreement with the guidelines of the Helsinki Declaration as revised in 1975. Permission to carry out the study was obtained from the institutional ethics committee. Informed consent was obtained from legal representatives of the participants, because all participants were unable to give consent. The measurements were performed in accordance with the guidelines of the Dutch Society for Doctors in the Care for people with an Intellectual Disability (NVAZ), which are outlined in a code called "Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans" [22]. The purpose of this code is to guide doctors in assessing resistance in persons with an intellectual disability. In line with this code, a participant's consistent distress or unhappiness was interpreted as a sign of lack of assent, and further participation in the study was reconsidered.

Design

The muscle tone and spasticity of 35 participants were measured twice (test and retest) with the MAS and MTS. The retest was conducted one week after the initial test. The participants were first assessed with the MAS and afterwards with the MTS. For each participant, both measurements were conducted at the same time of day and under the same conditions. Two observers performed the measurements. The interrater reliability of the MAS and MTS was determined from the measurements of the two observers. The test-retest reliability of the MAS and MTS was determined by using the test-retest measurements of observer 1.

Measures

Prior to the measurements, the observers and personal guides of the participants completed a checklist containing the exclusion criteria. Both observers were present at the time of the measurements. So that testing would not cause additional stress to the participants, we made sure that the participants were familiar with observer 2. We created a protocol describing how to administer the MAS and MTS based on the protocol of Gielen [18]. Gracies et al. [15] stated that training was associated with a highly significant improvement in reliability, so the observers were trained on how to perform the protocol. The training consisted of a brief explanation of the protocol and practical exercises. During the practical exercises, the results were compared and discussed. In the present study, we measured the most restricted joint motion of both the elbow and the knee.

Modified Ashworth Scale (MAS)

The MAS was carried out as follows. During five repetitions of a passive motion within one second, resistance was scored on the following 6-point scale [7]:

0 = No increased resistance
1 = Slightly increased resistance (catch followed by relaxation or minimal resistance at the end of the range of motion)
1 + = Slightly increased resistance (catch followed by minimal resistance throughout less than half of the range of motion)
2 = Clear resistance throughout most of the range of motion
3 = Strong resistance; passive movement is difficult
4 = Rigid flexion or extension

Catch is the phenomenon that suddenly a strong resistance occurs during a fast passive movement.

Modified Tardieu Scale (MTS)

The MTS consists of two measurements: R2 and R1 [16]. A goniometer was used for measuring the range of motion. The measurements were accurate to the 5-degree level. The R2 measurement consisted of slow motion performed within one second. The range of motion was measured with a goniometer. The R1 measurement consisted of fast motion performed within half a second. The range of motion immediately after the catch was measured with a goniometer.

Data analyses

The data were analysed using SPSS 15.0. The distribution of the data was determined and checked for normal distribution.

Feasibility

To assess feasibility, we compared the number of successful measurements per task to the total number of measurements. Since it only makes sense to use a test if a reasonable percentage of successful measurements can be made, this aspect of feasibility was considered to be sufficient if 85% of the measurements were successful [23, 24].

Test-retest reliability

Firstly, to determine whether significant differences between test and retest measurements exist, we analyzed the differences using the t-test or, in case of non-normally distributed data, the Wilcoxon signed rank test. The level of statistical significance was set at 0.05.

Secondly, intraclass correlation coefficients (ICC; two-way random, absolute agreement) of measurements 1 and 2 were computed. Reliability was considered to be acceptable if the ICC was greater than 0.75 and the 95% confidence interval (CI) was 0.3 or less. Reliability was considered to be very good if the ICC was greater than 0.9 [25].

Thirdly, Spearman correlation coefficients of measurements 1 and 2 were computed. Spearman's correlation was used because the data were not normally distributed. A correlation of 0.61 or more is considered good [26].

Fourthly, for the MTS, limits of agreement (LOA) between measurements 1 and 2 were calculated according to the procedure described by Bland and Altman [27]. LOAs were expressed together with the mean differences between measurements 1 and 2, and were judged whether they were narrow enough for the test to be of practical use, according to Atkinson and Nevill [28]. For the MAS, quadratically weighted kappa for measurements 1 and 2 was calculated. The quadratically weighted kappa is a measure of the proportion of agreement greater than that expected by chance. Values of kappa below .40 are generally considered to be clinically unacceptable, those within .41-.60 to be moderate, those within .61-.80 to be substantial, and those .81-1.00 to be almost perfect [29]. To obtain 95% confidence intervals (CIs) for the weighted kappa coefficients, we used the adjusted bootstrap percentile (BCa) method [30, 31] by employing the statistical programming language R [32].

Finally, the test-retest reliability of the MTS was considered reliable if (1) there were no significant differences between the test and retest measurements; (2) ICC was acceptable, as described above; (3) Spearman correlation coefficient was acceptable, as described above; and (4) LOA was acceptable, as described above. The test-retest reliability of the MAS was considered reliable if (1) there were no significant differences between the test and retest measurements; (2) ICC was acceptable, as described above; (3) Spearman correlation coefficient was acceptable, as described above; and (4) quadratically weighted kappa was almost perfect, with a 95% CI from substantial to almost perfect, as described above.

Interrater reliability

Firstly, to determine whether significant differences between the measurements of observers 1 and 2 exist, we analyzed the differences across measurements using a t-test or, in case of non-normally distributed data, the Wilcoxon signed rank test. The level of statistical significance was set at 0.05.

Secondly, the ICCs (two-way random, absolute agreement) of the measurements of observers 1 and 2 were computed. Reliability was considered to be acceptable if the ICC was greater than 0.75. Reliability was considered to be very good if the ICC was greater than 0.9 [25].

Thirdly, Spearman correlation coefficients of the measurements of observers 1 and 2 were computed. Spearman's correlation was used because the data were not normally distributed. A correlation of 0.61 or more is considered good [26].

Fourthly, for the MTS, LOAs between the measurements of observer 1 and 2 were calculated according to the procedure described by Bland and Altman [27]. LOAs were expressed together with the mean differences between the measurements of observers 1 and 2, and were judged whether they were narrow enough for the test to be of practical use, according to Atkinson and Nevill [28]. For the MAS, quadratically weighted kappa for the measurements of observers 1 and 2 was calculated. Values of kappa below .40 are generally considered to be clinically unacceptable, those within .41-.60 to be moderate, those within .61-.80 to be substantial, and those within .81-1.00 to be almost perfect [29]. We calculated 95% CIs for the weighted kappa coefficients (adjusted BCa method) [30, 31] by employing the statistical programming language R [32].

Finally, the interrater reliability of the MTS was considered acceptable if (1) there were no significant differences between the measurements of observers 1 and 2; (2) LOA was acceptable, as described above; (3) Spearman correlation coefficient was acceptable as described above; and (4) ICC was acceptable, as described above. The interrater reliability of the MAS was considered reliable if (1) there were no significant differences between the measurements of observers 1 and 2; (2) ICC was acceptable, as described above; (3) Spearman correlation coefficient was acceptable, as described above; and (4) quadratically weighted kappa was almost perfect, with a 95% CI from substantial to almost perfect, as described above.

Results

The data were not normally distributed; therefore, non-parametric tests were used to analyze the data. In all, 35 subjects participated in this study; 22 were male (62.9%), and 13 were female (37.1%). The mean age (SD) of the men was 35 (15) years, and that of the women was 31 (12) years. The characteristics of the study population are shown in Table 1.

Table 1. Characteristics of the study population.

Characteristics											
Intellectual disability		**Visual impairment**		**GMFCS level**		**Spasticity**		**Orthopedic defects**		**Epilepsy**	
Severe	22	Blind/severe	25	Level 4	11	Yes	17	Yes	30	Yes	32
Profound	13	Partially	10	Level 5	24	No	18	No	5	No	3

Feasibility

The percentages of successful measurements are shown in Table 2. Both the MAS and MTS showed a sufficient percentage of successful measurements.

Table 2. Percentages of successful measurements for the MTS and MAS.

Modified Tardieu Scale (MTS)	**Successful measurements week 1**	**Successful measurements week 2**
	Observer 1	Observer 1
Arm R2	100 %	97.1 %
Arm R1	100 %	97.1 %
Leg R2	97.1 %	94.3%
Leg R1	97.1 %	94.3%
Modified Ashworth Scale (MAS)		
Arm	100%	97.1 %
Leg	97.1 %	94.3%

Test-retest reliability of the MTS

Table 3 summarizes the statistical analyses for measurements 1 and 2 of the MTS.
There were no significant differences between measurements 1 and 2 ($p<0.05$). The ICC showed acceptable agreement for the R2 measurement of the arm and the R1 measurement of the leg. The ICC for the R1 measurement of the arm and the R2 measurement of the leg was not acceptable. The LOAs for both arm and leg measurements compared to median values were considerably large, which is clinically unacceptable. The Spearman correlation coefficient showed good correlation between the test-retest measurements for both arm and leg, except for the R2 measurement of the leg which was not acceptable.

Table 3. Summary of the statistical analyses for measurements 1 and 2 of the MTS for test-retest reliability.*

Modified Tardieu Scale (MTS)	Number of subjects	Median M1 (min-max)	Median M2 (min-max)	p level Wilcoxon	ICC	Mean difference ± LOA	Spearman coefficient
Arm R2	34	30 (0-95)	32.50 (0-95)	0.592	0.815	2.353 ± 35.2	0.792
Arm R1	34	55 (0-100)	57.50 (0-90)	0.890	0.627	6.029 ± 57.7	0.624
Leg R2	33	70 (0-135)	70 (0-120)	0.779	0.741	1.818 ± 44	0.402
Leg R1	33	67.50 (0-110)	75 (0-115)	0.089	0.850	-5.75 ± 29.8	0.680

*ICC, Intra Class Correlation Coefficient; LOA, Limits of agreement; m, measurement

Test-retest reliability of the MAS

Table 4 summarizes the statistical analyses for measurements 1 and 2 of the MAS.
There were no significant differences between measurements 1 and 2 ($p<0.05$). The ICC showed acceptable agreement between MAS measurements 1 and 2 for the arm and leg. The quadratically weighted kappa was almost perfect for the arm measurements and substantial for the leg measurements; the 95% CI values were substantial to almost perfect. The Spearman correlation coefficient demonstrated clear agreement between measurements 1 and 2 for the arm and leg.

Table 4. Summary of the statistical analyses for measurements 1 and 2 of the MAS for test-retest reliability.*

Modified Ashworth Scale (MAS)	Number of subjects	Median M1 (min-max)	Median M2 (min-max)	p level Wilcoxon	ICC	Quadratically weighted kappa	Spearman coefficient
Arm	35	1 0-4	1 0-4	0.430	0.853	0.82 .63-.92	0.808
Leg	33	1+ 0-4	1 0-4	0.771	0.813	0.78 .60-.89	0.757

*ICC, Intra Class Correlation Coefficient; m, measurement

Interrater reliability of the MTS

Table 5 summarizes the statistical analyses for the MTS measurements of observers 1 and 2. There were no significant differences between the measurements made by observers 1 and 2 ($p<0.05$). The ICC showed acceptable agreement between all the R2 and R1 measurements of observer 1 and 2. The LOAs of both arm and leg measurements compared to median values were considerably large, which is clinically not acceptable. The Spearman correlation coefficient between the R2 and R1 measurements of observer 1 and 2 were acceptable.

Table 5. Summary of the statistical analyses for the MTS measurements of observers 1 and 2.*

Modified Tardieu Scale (MTS)	Number of subjects	Median O1 (min-max)	Median O2 (min-max)	p level Wilcoxon	ICC	Mean difference ± LOA	Spearman coefficient
Arm R2	24	30 (0-95)	30 (0-110)	0.886	0.806	0.00 ± 40	0.813
Arm R1	24	55 (0-100)	70 (0-100)	0.540	0.851	5.870 ± 38	0.825
Leg R2	23	70 (0-135)	60 (0-115)	0.730	0.766	-0.91 ± 54	0.696
Leg R1	23	67.50 (0-110)	60 (0-120)	0.924	0.877	-0.45 ± 60	0.726

O, observer; ICC, intraclass correlation Coefficient; LOA, himits of Agreement.

Interrater reliability of the MAS

Table 6 summarizes the statistical analyses for MAS measurements of observers 1 and 2. There were no significant differences between the measurements taken by observers 1 and 2 ($p<0.05$). The ICC shows an acceptable agreement between the measurements of the arm and the leg of observers 1 and 2. The quadratically weighted kappa is almost perfect for the measurements of observers 1 and 2 of both the arm and the leg, the 95% CI are substantial to almost perfect. The Spearman correlation coefficient between the measurements of the arm and the leg of observers 1 and 2 is acceptable.

Table 6. Summary of the statistical analyses for the MAS measurements of observers 1 and 2. *

Modified Ashworth Scale (MAS)	Number of subjects	Median O1 (min-max)	Median O2 (min-max)	p level Wilcoxon	ICC	Quadratically weighted kappa	Spearman coefficient
Arm	23	1 0-4	1 0-4	0.350	0.894	.88 0.75-0.97	0.907
Leg	23	1+ 0-4	1+ 0-4	0.390	0.895	.88 .67-.96	0.858

* O, observer; ICC, Intraclass Correlation coefficient

Discussion

The purpose of our study was to determine the feasibility, the test-retest, and interrater reliability of the MAS and MTS in persons with PIMD. Our results demonstrated that the feasibility of the measurements was good, as an acceptable percentage of successful measurements was performed. The interrater reliability of the MAS was sufficient, with a substantial to almost perfect quadratically weighted kappa and an acceptable ICC. However, we found the interrater reliability of the MTS not to be clinically acceptable. Although the ICC indicated that the interrater reliability of the MTS was sufficient, the LOAs for both arm and leg measurements relative to median values were considerably large, which is clinically not acceptable. The MAS showed a sufficient test-retest reliability, with a substantial to almost perfect quadratically weighted kappa and an acceptable ICC . However, the test-retest reliability of the MTS was not sufficient due to its insufficient ICC, Spearman's correlations, and clinically unacceptable LOAs for both arm and leg measurements.

In our target group, the MAS showed a better test-retest and interrater reliability than the MTS, which contradicts the results of Gielen [18] and Mehrholz et al. [19]. In the study of Gielen [18], for the MAS the test-retest reliability calculated with Spearman's rho ranged from 0.66 to 0.81 (our study: 0.76-0.86) and the interrater reliability from 0.67 to 0.80 (our study: 0.86-0.91). For the MTS, Gielen's Spearman's rho was slightly better, with a range of 0.70 to 0.88 for intrarater reliability (our study: 0.40-0.79) and 0.70 to 0.82 for interrater reliability (our study: 0.70-0.83). As mentioned in the introduction, these contradictory findings also occurred in other target groups. All participants in our study population had impaired vision. Impaired vision may have contributed to differences in the faster movements of the MTS R2 measurements. Our subjects could not anticipate fast movements as well as their peers without visual impairments.

Compared to the study of Clopton et al. [13] in children with hypertonia, our ICC values for the MAS (0.81-0.85 for intrarater; 0.89-0.89 for interrater) were more sufficient than their ICC values (0.5-0.75 for intrarater; <0.5 for interrater). Mutlu et al. [10] performed a reliability study of the MAS in children with spastic cerebral palsy and found an intrarater reliability of 0.36-0.83 (ICC) and an interrater reliability of 0.54-0.78 (ICC), which is also less reliable than our ICC values for the MAS.

Our quadratically weighted kappa score for the interrater and intrarater reliability of the MAS was higher (interrater reliability 0.88; intrarater reliability 0.78-0.82) than the kappa (interrater reliability 0.51; intrarater reliability 0.59) of Ansari et al. [11], which tested the MAS in

persons with hemiplegia. In the study of Gregson et al. [9], which involved post-stroke patients, the kappa score for interrater reliability was 0.84 and for intrarater reliability was 0.83, which is comparable to the quadratically weighted kappa scores of the present study.

In our target group, the MTS had ICCs of 0.76-0.88 for interrater reliability and 0.63-0.85 for intrarater reliability, which is better than the interrater scores of Ansari et al. [17] in patients with hemiplegia (<0.56). In general, in the present study the interrater reliability of the MTS was better than the intrarater reliability. This difference may be due to changes in the condition of the participants at the time of the two measurements. This premise is also supported by the relatively large LOA outcomes, which indicated that there was too much variation at the individual level. The LOAs for both arm and leg measurements compared to median values was considerably large; most of the measurements departed more than 50% from the median. The LOAs were not narrow enough to indicate agreement between the two measurements of an individual. Therefore, we concluded that the LOAs for the MTS measurements were clinically unacceptable. Haugh et al. [14] suggested that the MTS is a more appropriate clinical measure of spasticity than the MAS. Therefore, we determined whether the LOA for MTS measurements of persons with spasticity was less than the LOA for measurements of persons without spasticity. However, the LOAs did not differ between these groups.

The purpose of our study was to determine the feasibility and reliability of the MAS and MTS in persons with PIMD. However, the validity of these instruments for measuring either spasticity or muscle tone was beyond the scope of this study. After the start of our study a manuscript was published entitled, "Stop using the Ashworth Scale for the assessment of spasticity" [34]. Taking their paper into account, we recommend that the validity of the MAS be examined further in future studies. Doing so, however, raises the issue of which specific parameter must be examined--the validity of measuring spasticity or the validity of measuring the quality of daily movements--since the MAS may be more suitable for measuring the latter. Ghotbi et al. [34] also obtained reliable measurements with the Modified Modified Ashworth Scale, supporting the recommendation that further research should be done on this instrument.

Given the outcomes of the ICC, quadratically weighted kappa, and the Spearman's correlation for the MAS compared to corresponding values obtained by other studies, the interrater and intrarater reliability of the MAS are sufficient in this target group. Furthermore, the usability of the MAS and MTS according to the observers appeared to be good. The protocol developed for the study functioned well for the observers. The observers were able to perform the tests properly after receiving training on the test protocol. The instructions were clear, and other physical therapists were able to follow the test protocol. The instruments also fit well within physical therapy. Moreover, testing was brief and the instruments were inexpensive.

Conclusions

Our research showed that the feasibility of using the MAS and MTS for measuring muscle tone in persons with PIMD is good. In our participants, the MAS showed sufficient test-retest and interrater reliability, whereas the MTS showed insufficient test-retest and interrater reliability. Therefore, the MAS may be a good method for evaluating the quality of daily movements in individuals with PIMD. Providing training and clear instructions on administering the MAS will improve reliability. Further research should aim to examine the validity of the MAS.

Recommendations

The feasibility of conducting MAS measurements in persons with PIMD is good. Adjustments in implementing the MAS are not necessary. Further research involving more participants may provide additional information.

Acknowledgements

This research was financed by Hanze University Groningen and Royal Dutch Visio De Brink. The authors kindly acknowledge and thank the participants for their participation in this study, their representatives for given permission to this and the physical therapists of Royal Dutch Visio The Brink, for their participation and for accompanying the participants during the tests. Furthermore, the authors kindly acknowledge W. Krijnen, PhD, for helping with statistical analyses.

References

1 Van der Putten A, Vlaskamp C, Reynders K, Nakken H. Children with profound intellectual and multiple disabilities: the effects of functional movement activities. Clinic Rehabil. 2005;19, 613-620.

2 Palisano R, Hanna SE, Rosenbaum PL, Rusell DJ, Walter SD, Wood EP, Raina PS, Galuppi BE. Validation of a model of Gross Motor Function for Children With Cerebral Palsy. Physical Therapy. 2000;80:974-985.

3 Shevell MD, Dagenais L, Hall N. Comorbidities in cerebral palsy and their relationship to neurologic subtype and GMFCS level. Neurology. 2009;72:2090-2096.

4 Lance JW. What is spasticity? The Lancet. 1990;335:606.

5 Stedman TL. Stedman's Medical Dictionary, 27th Edition. 1999, Philadelphia: Lippincott, Williams and Wilkins Company.

6 Evenhuis HM, Sjoukes L, Koot HM, Kooijman AC. Does visual impairment lead to additional disability in adults with intellectual disabilities? J Intellect Disabil Res. 2009;53:19-28.

7 Bohannon RW & Smith MB. Interrater Reliability of a Modified Ashworth Scale of Muscle Spasticity. Physical Therapy. 1987;67:206-207.

8 Pandyan AD, Johnson GR, Price CI, Curless RH, Barnes MP, Rodgers HA. Review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. ClinicRehabil. 1999;13:373-383.

9 Gregson JM, Leathley M, Moore AP, Sharma AK, Smith TL, & Watkins CL. Reliability of the Tone Assessment Scale and the Modified Ashworth Scale as clinical tools for assessing poststroke spasticity. Arch Phys Med Rehab.1999;80:1013-1016.

10 Mutlu A, Livanelioglu A, Gunel MK. Reliability of Ashworth and Modified Ashworth scales in children with spastic cerebral palsy. BMC Musculoskel Dis. 2008;10:9-44.

11 Ansari NN, Naghdi S, Arab TK, Jalaie S. The interrater and intrarater reliability of the Modified Ashworth Scale in the assessment of muscle spasticity: limb and muscle group effect. NeuroRehabilitation. 2008;23:231-237.

12 Yam WK & Leung MS. Interrater reliability of Modified Ashworth Scale and Modified Tardieu Scale in children with spastic cerebral palsy. J Child Neurol. 2006;21:1031-1035.

13 Clopton N, Dutton J, Featherston T, Grigsby A, Mobley J, Melvin J. Interrater and intrarater reliability of the Modified Ashworth Scale in children with hypertonia. Pediatric physical therapy. 2005;17:268-274.

14 Haugh AB, Pandyan AD, Johnson GR. A systematic review of the Tardieu Scale for the measurement of spasticity. Dis Rehabil. 2006;28:899-907.

15 Gracies JM, Burke K, Clegg NJ, Browne R, Rushing C, Fehlings D, Matthews D, Tilton A, Delgado MR. Reliability of the Tardieu Scale for assessing spasticity in children with cerebral palsy. Arch Phys Med Rehabil. 2010;91:421-428.

16 Mackey AH, Walt SE, Lobb G, Stott NS. Intraobserver reliability of the Modified Tardieu Scale in the limb of children with Hemiplegia. Dev Med Child Neurol. 2004;46:267-272.

17 Ansari NN, Naghdi S, Hasson S, Azarsa MH, Azarnia S. The Modified Tardieu Scale for the measurement of elbow flexor spasticity in adult patients with hemiplegia. Brain Injury. 2008;22:1007-1012.

18 Gielen EJJM. Is spasticiteit te meten bij mensen met een verstandelijke beperking? Een betrouwbaarheids- en validiteitsonderzoek van de Modified Ashworth Scale en de Tardieu Schaal bij mensen met een verstandelijke beperking. Scriptie in het kader van de cursus Scholing in Wetenschap III; Amersfoort, Nederlands Paramedisch Instituut, 2005.
19 Mehrholz J, Wagner K, Meiβner D, Grundmann K, Zange C, Koch R, Pohl M. Reliability of the Modified Tardieu Scale and the Modified Ashworth Scale in adults patients with severe brain injury: a comparison study. Clinic Rehabil. 2005;19:751.
20 World Health Organization (1992). International statistical classification of diseases and related health problems: tenth revision (ICD-10), 369–370.
21 World Health Organization (WHO). International Classification of Functioning, Disability and Health. Geneva, 2001.
22 Dutch Society for Doctors in the Care for people with an Intellectual Disability (NVAZ) Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans, behavioural code for doctors in the assessment of resistance among people with an intellectual disability, 1999.
23 Lemmink KAPM. De Groninger Fitheidstest voor Ouderen. Ontwikkeling van een meetinstrument. 1996, RUG.
24 Malmberg JJ, Miilunpalo SI, Vuori IM, Pasanen ME, Oja P, Haapanen-Niemi NA. A health-related fitness and functional performance test battery for middle-aged and older adults: feasibility and health-related content validity. Arch Phys Med. 2002;83(5):666-677.
25 Portney LG, Watkins MP. Foundations of clinical research. 2nd ed. Upper Saddle River, 2000, Prentice-Hall Inc, New Jersey.
26 Feinstein AR. Clinimatrics. 1987, New Haven, Yale University Press, 184-185.
27 Bland JM & Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986;1:307-310.
28 Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26(4):217-38.
29 Sim J & Wright CC. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. Physical Therapy. 2005;85;257-268.
30 Efron B. Better bootstrap confidence intervals. J Am Statist Ass. 1987;82:171-185.
31 Davison AC & Hinkley DV. Bootstrap Methods and Their Application. 1997, Cambridge University Press.
32 R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2009, Vienna, Austria. URL http://www.R-project.org.
33 Fleuren JF, Voerman GE, Erren-Wolters CV, Snoek GJ, Rietman JS, Hermens HJ, Nene AV. Stop using the Ashworth Scale for the assessment of spasticity. J Neurol Neurosurg Psych. 2010;81:46-52.
34 Ghotbi N, Ansari NN, Naghdi S, Hasson S, Jamshidpour B, Amiri S, Inter-rater reliability of the Modified Ashworth Scale in assessing lower limb muscle spasticity. Brain Injury. 2009;23:815-819.

Chapter 8

Heart Rate Pattern as an Indicator of Physical activity in persons with Profound Intellectual, and Multiple Disabilities.

A. Waninge
A.A.J. van der Putten
R.E. Stewart
B.Steenbergen
R. van Wijck
C.P. van der Schans

Submitted

Abstract

Background As physical fitness is related to physical activity, it is important to gain insight into the physical activity levels of persons with Profound Intellectual and Multiple Disabilities (PIMD). Heart rate monitoring may be used as an indicator of activity levels, yet a correct method for dating heart rate patterns of subjects with PIMD has so far not been researched.
Objective The purpose of this study was twofold. First, this study examines the activity levels of persons with Profound Intellectual and Multiple Disabilities (PIMD) based on their heart rate patterns, contrasted against American College of Sports Medicine (ACSM) guidelines of healthy physical activity. Second, this study describes the relation between various covariates and heart rate patterns and proposes adherent classification.
Method Using a heart rate monitor, heart rate patterns of 24 subjects with PIMD were measured during 6 days. Heart rate intensity was calculated using heart rate reserves. Physical activity levels were also measured with questionnaires. Data were analyzed using multilevel analysis.
Results The results show that the mean heart rate zone of the participants over six days is 3.196, indicating 32 % of the heart rate reserve. The intensity ranged over a heart rate reserve from 1 to 62 %. Wide ranges in heart rate between participants and within one day have been shown. However, between days we found small ranges in heart rate. Heart rate monitoring is a reliable measurement for measuring physical activity. Participants could be classified in 4 classes according to heart rate. Time of day, physical activity and age have significant influence on heart rate.
Conclusions In conclusion, persons with PIMD are not sufficiently physically active based on the guidelines of ACSM. Heart rate monitoring seems to be a reliable indicator of physical activity, but exploration of the other influential factors, such as emotions and personal factors, is recommended.

Introduction

It is important to gain insight into the physical activity levels of persons with profound intellectual, visual, and severe motor disabilities (profound intellectual and multiple disabilities, PIMD). Persons with PIMD risk low levels of physical activity [1, 2] and the associated negative effects on health. Physical activity improves both mental health, physical health, physical fitness [3], and participation in daily life [4]. Physical activity is defined as any bodily movement produced by skeletal muscles that results in energy expenditure [5]. However, for a substantial gain in physical fitness, the ACSM guidelines state that physical activity has to be performed 5 days a week for at least 30 minutes a day with an intensity of more than 55 % of the heart rate reserve [6].

Persons with intellectual disabilities are often not sufficiently active to achieve benefits in health or improve fitness levels [2, 7, 8]. Additionally, physical fitness in persons with a visual disability is poorer than in persons without disabilities [9, 10, 11], and persons with both severe intellectual and visual disabilities have a high chance of experiencing a variety of limitations in daily functioning such as inactivity, insecure movement and little initiative [12]. However, research into the physical activity levels in persons with PIMD is limited and knowledge on the topic is scarce.

Due to severe multiple disabilities, physical activity levels of persons with PIMD are difficult to reliably quantify [13]. These persons often also suffer from sensory integration problems [14], and inadequacies in perception and motor-reproduction [15, 16]. Furthermore, co-morbidity, such as cerebral palsy, is more frequent in those with intellectual disabilities than in the general population [17]. As a result, persons with PIMD have generally very limited mobility, use a wheelchair [18], and have a Gross Motor Function Classification System level IV and V (GMFCS) [19]. Consequently, normal tests, such as the usage of walking, are not applicable for persons with PIMD [13]. Moreover, the presumed low levels of activity in persons with such profound disabilities are often not accurately presented by relatively insensitive measurement devices, like activity monitors [13]. In addition, there are no existing algorithms for predicting activity energy expenditure of persons with PIMD.

Physical activity studies often use a combination of assessment methods [20, 21] including heart rate monitoring, which is an objective method [22], combined with direct observation, which is a criterion method [22]. Heart rate monitoring may be used as an indicator of activity levels when assuming a relationship between activity intensity and heart rate [23, 24]. Heart rate monitoring appears to be sufficiently valid to use in creating broad physical activity categories (e.g. highly active, somewhat active, sedentary) [25]. As stated before, only heart rates of more than 55% of the heart rate reserve may gain profit for physical fitness, if obtained during 5 days in a week [6]. Thus, heart rate monitoring may tell us if we can actually increase the persons' fitness.

However, the correct method of dating heart rate patterns of individuals with PIMD as well as the correlation between heart rate monitoring and activity levels for this specific group have so far not been subject to research. What is more, also the influence of covariates such as gender and age on the heart rate patterns of individuals with PIMD is unknown. For persons with and without disabilities, physical activity is gender related [18]. Heart rate is related to age [26], gender and activity [27]. As persons with PIMD often suffer from co-morbidity such as motor disability, spasticity and sensory disabilities, it seems useful to examine the influence of these covariates on heart rate height as well.

Furthermore, practical experience learns that as a consequence of co-morbidity, the skills of persons with PIMD vary greatly. Moreover, Vlaskamp et al. [28] found days in the PIMD activity centres to be highly structured, with each activity taking place at the same time and day. These findings suggest the possibility of a relation between heart rate patterns and subgroups plus time of day.

The purpose of this study, therefore, was fourfold: firstly, to determine the activity levels of persons with PIMD based on heart rate patterns when compared to ACSM guidelines of healthy physical activity; secondly, to analyze heart rate patterns according to group differences, days, time of day and to establish adherent classification in heart rate height and patterns; thirdly, to determine the relation between heart rate patterns and observed level of activity in persons with PIMD; and, fourthly, to examine the influence of covariates such as gender, age, and common co-morbidity (motor disabilities, spasticity and sensory disabilities) on heart rate patterns.

Materials and methods

Participants

The target population of our study comprises of persons with PIMD, characterized by severe or profound intellectual disability indicated by an intelligence quotient under 40 points. The participants have a developmental level lower than six years (International Association for the Scientific Study of Intellectual Disabilities, IASSID) [29], and are thus severely limited in self-care, continence, communication, and mobility [30].

All participants were recruited from a Dutch residential care facility, which houses 200 persons with severe or profound intellectual and visual disabilities. The inclusion criteria were: presence of severe or profound intellectual disability, visual disability, and motor disability with GMFCS level IV or V [19]. For 48 persons, representatives were requested to give a written permission for participation in this study, of which 30 were obtained. Both a physician specialised in mental disabilities and a behaviour scholar screened the participants for our exclusion criteria, being severe psychological problems or somatic diseases defined as chronic diseases and/or diseases that do not resolve in the short term.

Four persons were excluded from the study because they showed one of these problems or diseases. The exclusion criteria at the time of the measurements were: general illness or fever; taking antibiotics; worsening of asthma, epilepsy (recent insult or epileptic fits), fresh wound(s)/bruise(s), or other factors causing pain during movement; or stress due to the subject's behavior just before the measurement dates. Two persons were excluded because they presented one of these signals. Figure 1 presents the sampling scheme of persons included in the study.

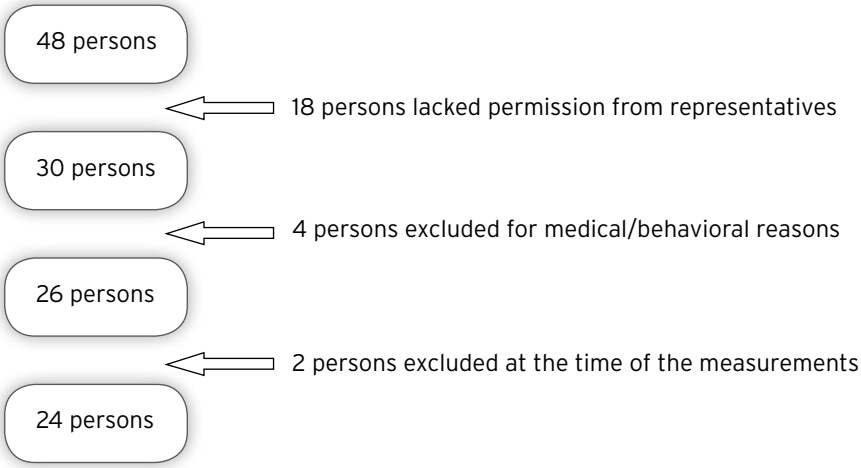


Figure 1. Sampling scheme of subjects included in the study

All participants were classified according to the GMFCS [19]. Furthermore, visual and auditive impairments of the participants were classified according to WHO guidelines [4]: a distinction was made between being severely partially sighted and being partially sighted as well as between severe hearing loss, slight hearing loss and normal hearing. Spasticity was classified as unilateral, bilateral or unknown [31]. Orthopedic defects were used as an indicator of locomotor disabilities and classified as present or not present.

Ethical statement

This study was performed in agreement with the guidelines of the Helsinki Declaration as revised in 1975. Permission to carry out the study was obtained from the institutional ethics committee. Informed consent was obtained from representatives of the participants, because the participants were not able to give consent. The measurements were performed in accordance with the behavioural code section entitled 'Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans' [32]. Consistent distress or unhappiness was interpreted as a sign of lack of assent and further participation in the study was reconsidered.

Study design

Heart rate patterns were measured in each participant 8 hours a day for a period of six days. Every 15 minutes measurements were conducted, resulting in a total amount of measurements a day of 32 (8 hours, 4 times 15 minutes). 5 Out of the 6 test days were weekly, the remaining day fell in the weekend. Parallel with heart rate measurements, physical activities were registered using direct observation, noted down in score lists.

Measures

Heart rate patterns were measured with a heart rate monitor (Polar RS 800, Kempele, Finland) whose heartbeat data were transferred later to a computer. Heart rate was monitored in every participant during 6 days during 8 hours a day.

Data regarding physical activity were registered with the use of a questionnaire, which was filled out by both personal caregivers at the living group as support staff of the activity centre. Physical activity was coded as 'Targeted physical activity Yes' or 'Targeted physical activity No'. Moving with the wheelchair inside or outside, transfer, active sitting without support, gymnastics with a gymnastic instructor, physical therapy, playing with a ball, and 'dancing' on music were all examples of 'Targeted physical activity Yes'. Listening to music, watching television or lying down on a bed were all examples of ''Targeted physical activity No'.

Data analysis

Heart rate zones

Peak heart rate, rest heart rate and heart rate reserves differ for each person, which makes them difficult to compare. By calculating heart rate zones according to the equation of Karvonen [33], it is possible to compare the zones of the participants with each other.

Heart rate zones are calculated as follows. First, each participant's peak heart rate was estimated using the formula of Fernhall [34] for participants with intellectual disabilities: 210 – 0.56 (age) – 15.5. Due to the motor disabilities of the participants, no other non-invasive measure could be performed. Secondly, the participants resting heart rate was determined by taking the median of fifteen morning heart rate measurements. Thirdly, using the participants resting heart rate, the heart rate reserve was calculated by subtracting resting heart rate of estimated peak heart rate. Finally, the heart rate reserve was divided in 10 zones, each zone consisting of 10% of the heart rate reserve. The heart rate of a participant during each 15 minutes was classified in these zones. For instance: resting heart rate of 50 beats per minute (bpm), maximum heart rate of 180 bpm; heart rate reserve is 130 bpm; each heart rate zone exists of 13 heart rates, the first zone is from 50 to 63; the second from 63 to 76; and so on (Table 1).

Table 1. Example of the heart rate zones for healthy persons

Heart Rate	Zone	Activity	Percentage of heart rate reserve (HRV)
1	50-63	Rest	1-10
2	63-77		10-20
3	77-90	Quiet moving	20-30
4	90-103		30-40
5	103-116	Moderately intensive activity	40-50
6	116-129	ACSM guideline of healthy physical activity	50-60
7	129-142	Intensive activity	60-70
8	142-155		70-80
9	155-168	Very intensive activity	80-90
10	168-180		90-100

Patterns and classes in heart rate

Heart rate of the 24 participants was measured eight hours a day during six days, with measurements being conducted every 15 minutes. In order to determine the activity levels of persons with PIMD compared with ACSM guidelines of healthy physical activity, an overview of the heart rate zones is presented, along with a day to day outline of the mean prevalence of heart rate zones of the participants.

Furthermore, the mean and the ranges of heart rate zones of the participants are calculated. To gain insight into the heart rate patterns of persons with PIMD, three decomposing variance components are involved, using linear mixed model: 1) between persons, 2) within persons between days, and 3) within days. The mean heart rate and variance proportion component (VPC) were calculated between persons, within persons and between days. The VPC as an indicator of variance in heart rate zone, is calculated by dividing variance by the total variance. Using the variance proportion component the generalizability coefficient for the relative differences [35] was calculated. A generalizability coefficient of 0.80 or more indicates a sufficient reliability. Furthermore, in order to identify distinct groups of heart rate patterns and to examine these classes in heart rate, we used a latent class analysis [36, 37]. As the dependent variable was a count variable, a Poisson distribution was used for this analysis.

Relations between heart rate patterns and level of activity

To determine how heart rate relates to the level of physical activity, we estimated equations of physical activity as a dependent variable of heart rate. Furthermore, we examined the influence of 'time of the day' on this relation.

Influence of covariates

The influence of the covariates gender, age, time of day, daily activities, motor disabilities, spasticity, and sensory disabilities were evaluated in the mixed model.

Results

The data were analysed using SPSS 16.0 and multilevel analysis with the computerprogram MIwin [38].

In total, 24 persons with PIMD participated in this study. Six women participated with a mean age (SD) of 30 years (17), the mean age (SD) of the men was 36 years (15). Table 2 shows the characteristics of the participants.

Table 2. Characteristics of the participants

		Gender	
		Men	**Women**
Intellectual disability	Severe	9	3
	Profound	9	3
	Total	18	6
Visual impairments	Blind/Severe	13	3
	Partially	5	3
	Total	18	6
Orthopedic defects	Yes	16	6
	No	2	0
	Total	18	6
Spasticity	Yes	16	6
	No	2	0
	Total	18	6
GMFCS level	IV	5	1
	V	13	5
	Total	18	6

Patterns and classes in heart rate

Table 3 shows a day to day outline of the mean prevalence of heart rate zones in persons with PIMD, indicating that the participants reach no heart rates more than 55% of their heart rate reserves for a consecutive 30 minutes.

Table 3. Day-to-day outline of the mean prevalence of heart rate zones in persons with PIMD

Percentage of time in zone						
Percentages of Heart rate reserve	**Day 1**	**Day 2**	**Day 3**	**Day 4**	**Day 5**	**Day 6**
1-10%	9%	7%	6%	10%	1%	6%
11-20%	28%	26%	26%	23%	29%	23%
21-30%	26%	26%	24%	18%	32%	30%
31-40%	25%	29%	23%	28%	27%	28%
41-50%	11%	10%	15%	18%	10%	11%
51-60%	1%	2%	4%	3%	1%	2%
61-70%			2%			
71-80%						
81-90%						
91-100%						
Total time	100%	100%	100%	100%	100%	100%

The mean heart rate zone over the six test days is shown under the heading 'intercept' model 0 in table 4, and is 3.196, indicating 32 % of the heart rate reserve. The results range from 1 to 62 % of the heart rate reserve (see table 3).

The model in table 4 shows the variance between participants (0.911), between days (0.161) and within one day (0.429). As the total variance is the sum of these variances (1.501), the VPC between the participants is 60.7%, between days 10.7%, and between different times of the day 28.6%. These VPC's indicate that the variation between the test days is relatively low, suggesting that future research can do with one test day instead of six. The generalizability coefficient for the relative differences, calculated with the variance proportion components, is 0.85, indicating that heart rate monitoring is a reliable measurement.

Table 4. Variability of heart rate and linear mixed models.

Fixed	**Model**	
(person level)	B-value	SE
Intercept	4.217	0.655
Time of day	-0.007	0.001
Activity Targeted activity vs no targeted (ref)	0.021	0.021
Age	-0.023	0.012
Gender Male vs female (ref)	-0.041	0.409
Orthopaedic defects No orth defects vs orth defects (ref)	-0.257	0.577
Spasticity Bilateral vs unilateral parese (ref)	0.499	0.867
Visual impair Severe partially vs partially impaired vision (ref)	0.185	0.39
Auditive impairment Severe hearing loss vs slight hearing loss vs normal hearing	-0.606 -0.735	0.593 0.462
Random		
Participant	0.911	
Days	0.161	
Time of day	0.429	
Total var	1.501	
2Loglikelihood	10302	

B-value, the regression-coefficient; SE, standard error; vs, versus; ref, reference value.

With the variable age, time of the day and the observations measured at six different days, we found a four class solution that gives a clear clinical interpretation [36, 37]. The four class classes were: high heart rate zone (class 1), middle stable heart rate zone (class 2), low heart rate zone (class 3) and variation in the middle heart rate zone (class 4).

Relation between heart rate patterns and level of activity

The relation between time of the day and heart rate is significant (regression-coefficient-0.007 with SE 0.001; p< 0.01) (see model in table 4). The relation between physical activity and heart rate shows a regression-coefficient of 0.041 with SE of 0.021 (p= 0.02), which is a significant relation too. However, if 'time of the day' is brought into the equation, there is no significant influence of 'physical activity' on heart rate anymore (regression-coefficient 0.022 with SE 0.021; p=0.147), only 'time of the day' remains of influence (-0.007 with SE 0.001; p<0.01).

Therefore, we examined the relation between 'time of day' and 'daily activity', based on the questionnaires filled out by the caregivers. Figure 2 shows the percentage active persons as a function of the time of the day. Figure 2 illustrates that most participants are active between 9.00 and 12.00 o'clock. During the day the percentage active persons decreases. Between 12.00 and 13.00 o'clock most participants seem to rest. Between 13.30 and 16.00 o'clock most participants become active again, but afterwards activity decreases once more.

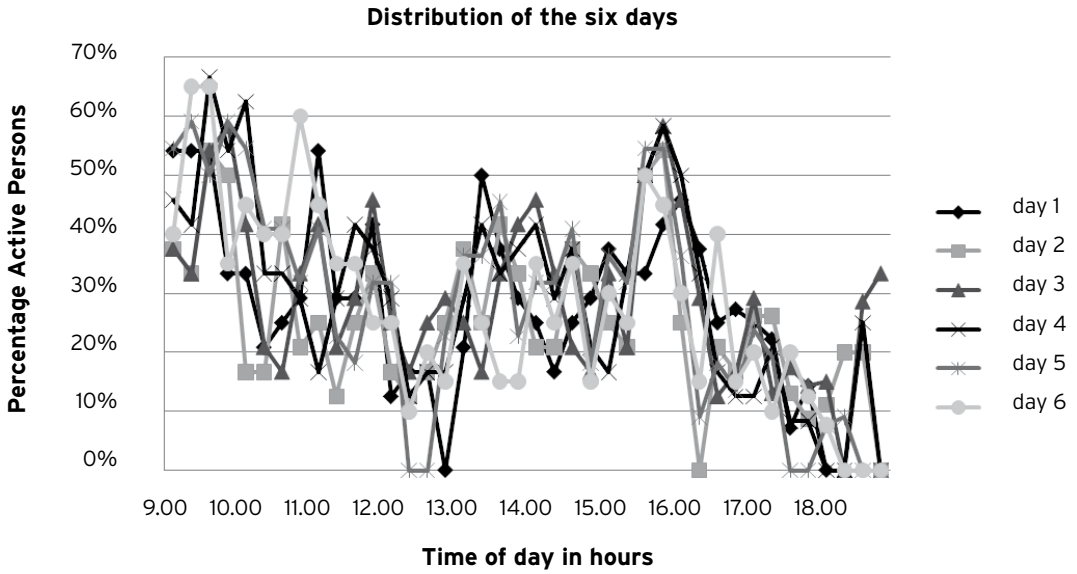


Figure 2. Relation between 'time of the day' and 'activity': the percentage active persons as a function of the time of the day, based on the questionnaires.

Influence of covariates

Table 4 shows the influence of covariates. Age has a significant influence. Higher ages yielded lower heart rate zones. Also time on the day has a statistical significant influence: later on the day yielded lower heart rate zones. Gender, spasticity, intellectual level and hearing had no significant influence.

Discussion

This study examined firstly the activity levels of participants with PIMD based on their heart rate patterns when contrasted against ACSM guidelines of healthy physical activity [6]. Secondly, this study researched the relation between various covariates and heart rate patterns and proposed adherent classification.

Our study shows that the participants reach no heart rates more than 55% of their heart rate reserves for a consecutive 30 minutes, which indicates that they are not sufficient physical active based on the guidelines of ACSM [6]. Wide ranges in heart rate between participants have been shown. In addition, within one day wide ranges in heart rate are present. However, between days we found small ranges in heart rate. The generalizability coefficient (>0.80) indicates that heart rate monitoring is a reliable method for measuring physical activity. Participants could be classified in 4 groups according to heart rate. Time of day, physical activity and age have a significant influence on heart rate. However, the relation between time of day and physical activity is very strong and when corrected for 'time of day', 'physical activity' ceases to have a significant influence on heart rate. As a consequence, heart rate seems to be an indicator of physical activity of individuals with PIMD, but further exploration of other possible factors is still needed.

This study is an important first step in exploring heart rate patterns of the intellectually disabled for a full day. Due to their locomotor disabilities, examination of the peak heart rate with a maximal heart rate test is not possible. Therefore, we had to use the formula of Fernhall [34], which is an estimate and may therefore be less reliable. Based on our results we recommend further exploration of heart rate patterns in this target group.

The mean heart rate zone of all participants over six days is 3.2. This indicates they use relatively little, only 20 to 30 percent, of their heart rate reserves. Although the range of percentages of heart rate reserves was found to be between 1 and 62 %, levels higher than 55 % for at least a consecutive 30 minutes were not reached. It appeared to be difficult to activate these persons into physical activity. This may be explained by the multiple disabilities of our target group. Persons with GMFCS levels IV and V have little skills to move actively and are hindered by visual impairment and lack of comprehension. The individual combinations of these limitations may explain the four classes we distinguished in heart rate patterns. Future research should be directed towards the examination of these classes and the possible influence of interventions on these classes.

Caregivers in the living situation as well as caregivers in the activity centre filled out the questionnaires for registration of physical activity used in our analysis. Analysis of this registration brings forward a figure similar to that of the heart rate patterns. The mean number of offered motor activities was 0.8 per day (range: 0-3.3; SD: 1.1) and the mean duration of motor activation was 26 minutes per day (range: 0-163; SD: 35) [39]. The present study yielded a negative correlation between heart rate and age. Van der Putten & Vlaskamp [39] found a negative correlation between daily activity and age. As heart rate height and physical activity are supposed to be related [3], our results are in line with the results of Van der Putten & Vlaskamp [39].

Furthermore, the majority of observed motor activation took place during 'daily care situations', whereas only 18% of these situations existed of targeted motor activation. 61 % Of the activities were passive in nature [40]. The conclusions of these authors were that motor activation is a minor part of the support of persons with PIMD. Moreover, analysing the heart rate

patterns in this study we can conclude that they do not meet the ACSM recommendations for health related physical fitness [6].

The significant relation between time of the day and heart rate may partly be declared by the circadian rhythms. The circadian system, driven by the suprachiasmatic nucleus, regulates properties of cardiovascular function, like blood pressure and heart rate [41, 42].

Furthermore, the relation between time of day and heart rate may be an illustration of the highly structured days in the activity centres of our participants. This is in line with the findings of Vlaskamp et al [28]: in 6 care facilities in the Netherlands it seemed to be common practice to have one activity in the morning, mostly between 10 and 11 AM, and one in the afternoon, roughly between 2 and 3 PM. This degree of structure already made us expect a low difference in heart rate patterns between days. Yet, is such structure desirable? Would an alternating day rhythm result in more activation of persons with PIMD? On the other hand, the advantage of the low difference in test results between days is that for future research one day of testing suffices, which is efficient.

We found no significant relation between heart rate and gender, intellectual level or co-morbidity as motor disabilities, spasticity and sensory disabilities when looking at 8-hours heart rate patterns. In order to test for any possible significant relation future research could set up subcategories, controlling for groups with a specific co-morbidity using groups without that specific co-morbidity.

Heart rate height as a predictor of antisocial behavior in adolescents was examined in a study of Sijtsema et al [43]. Their findings showed that heart rate measures obtained with a strict acquisition and analysis protocol were associated with antisocial behavior in boys but not in girls. In other studies it was found that in childhood and adolescence, low heart rate is one of the strongest correlates of antisocial behavior [44]. Furthermore, according to the stimulation-seeking theory, some adolescents are constantly under aroused, which is presumably marked by a low heart rate and a subjective unpleasant state [45]. Given the low heart rate zones in our study population, aforementioned research suggests examining the subjective well-being of these participants may be of crucial importance.

Physiological outcome measures as an indicator of subjective well-being were already explored by Vos et al [46] in persons with PIMD. People with PIMD showed more parasympathic activation when experiencing negative emotions. Most likely this is due to attention regulating processes. They also show a higher heart rate when the emotion intensifies. Nevertheless, the authors caused the readers to be careful when interpreting their findings since there were several limitations to their research. Taking all these findings into account, further research on the relation between heart rate of persons with PIMD and psychosocial variables such as emotions is recommended.

In conclusion, persons with PIMD do not attain sufficient activity levels according to ACSM guidelines. Heart rate monitoring seems to be a reliable indicator of physical activity. Time of day and age have considerable influence on heart rate patterns. However, the observed classes in heart rate patterns suggest other, probably more personal and psychosocial factors to have a significant influence on heart rate patterns as well. Further research into these factors is recommended.

Acknowledgements

This research was financed by Hanze University Groningen and Royal Dutch Visio De Brink. The authors kindly acknowledge and thank the participants for their participation in this study, their representatives for given permission to this and both personal caregivers at the living group and support staff of the activity centre of Royal Dutch Visio The Brink, for filling out the questionnaires. Furthermore, the authors kindly acknowledge and thank Floor J. Hettinga, PhD, for advice on physiological topics.

References

1 Emerson E. Underweight, obesity and exercise among adults with intellectual disabilities in supported accommodation in Northern England. J Intellect Disabil Res. 2005;49;134-43.

2 Lahtinen U, Rintala P, Malin A. Physical performance of individuals with intellectual disability: a 30 year follow up. Adapt Phys Act Q. 2007;24: 125-143.

3 Bouchard C, Shepard RJ, Stephens T. Physical activity, Fitness and Health. 1994, Human Kinetics Publishers, ChampaignIL.

4 World Health Organization (WHO). International Classification of Functioning, Disability and Health. Geneva, 2001.

5 Capsersen CJ, Powell KE, Christenson GM. Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research. Public Health Reports. 1985;100(2);127-131.

6 American College of Sports Medicine Position Stand. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. Med Sci Sports Exerc. 1998;30: 975-91.

7 Frey GC & Chow B. Relationship between BMI, physical fitness, and motor skills in youth with mild intellectual disabilities. Int J Obes. 2006;30:861-867.

8 Temple VA, Frey GC, Stanish HI. Physical activity of adults with mental retardation: review and research needs. Am J Health Promot. 2006;21:2-12.

9 Häkkinen A, Holopainen E, Kautiainen H, Sillanpää E, Häkkinen K. Neuromuscular function and balance of prepubertal and pubertal blind and sighted boys. Acta Paediatr. 2006;95(10):1277-83.

10 Hopkins WG, Gaeta H, Thomas AC, Hill PM. Physical fitness of blind and sighted children. Eur J of Appl Phys Occ Phys. 1987;56: 69-73.

11 Seemungal BM, Glasauwer S, Gresty MA, Bronstein AM. Vestibular perception and navigation in the congenitally blind. J of Neurophys. 2007;97(6):4341-56.

12 Evenhuis HM, Sjoukes L, Koot HM, Kooijman AC. Does visual impairment lead to additional disability in adults with intellectual disabilities? J Intellect Disabil Res. 2009;53(1):19-28.

13 Warms C. Physical Activity Measurement in Persons with chronic and disabling conditions, methods, strategies and issues. Fam Community Health. 2006;29:788-799.

14 Carmeli E, Bar-Yossef T, Ariav C, Paz R, Sabbaq H, Levy R. Sensorimotor impairments and strategies in adults with intellectual disabilities. Motor Control. 2008;12(4):348-61.

15 Reimer AM, Cox RF, Boonstra NF, Smits-Engelsman BC. Effect of visual impairment on goal-directed aiming movements in children. Dev Med Child Neur. 2008;50(10):778-83.

16 Shinkfield AJ, Sparrow WA, Day RH. Visual discrimination and motor reproduction of movement by individuals with mental retardation. Am J Ment Retard. 1997;102(2):172-81.

17 Van Schrojenstein Lantman-de Valk HMJ, Metsemakers JFM, Haveman MJ, Crebolder HFJM. Health problems in people with intellectual disability in general practice: A comparative study. Family Practice. 2000;17:405–407

18 Putten AAJ van der, Vlaskamp C, Reynders K, Nakken H. Children with Profound Intellectual and Multiple Disabilities: the effects of functional movement activities. Clin Rehabil. 2005;19:613-620.

19 Palisano R, Hanna SE, Rosenbaum PL, Rusell DJ, Walter SD, Wood EP, Raina PS, Galuppi BE. Validation of a model of Gross Motor Function for Children With Cerebral Palsy. Phys Therapy. 2000;80:974-985.

20 Guinhouya BC, Apété GK, Hubert H. Evaluation of habitual physical activity of children during clinical and epidemiological trials. Sante Publique. 2009;21(5):465-78.
21 Bock De F, Menze J, Becker S, Litaker D, Fischer J, Seidel I. Combining Accelerometry and Heart Rate for Assessing Preschoolers' Physical Activity. Med Sci Sports Exerc. 2010;42(12):2237-43.
22 Vanhees L, Lefevre J, Philippaerts R, Martens M, Huygens W, Troosters T, Beunen G. How to assess physical activity? How to assess physical fitness? Eur J Cardiovasc Prev Rehabil. 2005;12(2):102-14
23 Hunter J. Energy costs of wheelchair propulsion by elderly and disabled people. Int J Rehabil Res. 1987;50-54.
24 Littlewood RA, Davies PSW, Cleghorn GJ, Grote RH. Physical activity cost in children following an acquired brain injury—a comparative study. Clin Nutrition. 2004;23;99-104.
25 Sirard JR, Pate RR. Physical activity assessment in children and adolescents. Sports Med. 2001;31(6):439-54.
26 Baynard T, Pitetti KH, Guerra M, Unnithan VB, Fernhall B. Age-related changes in aerobic capacity in individuals with mental retardation: a 20-yr review. Med Sci Sports Exerc. 2008;40(11):1984-9.
27 Climstein M, Pitetti KH, Barrett PJ, Campbell KD. The accuracy of predicting treadmill VO2max for adults with mental retardation, with and without Down's syndrome, using ACSM gender- and activity-specific regression equations. J Intellect Disabil Res. 1993;37:521-3.
28 Vlaskamp C, Hiemstra SJ, Wiersma LA & Zijlstra BJH. Extent, duration, and content of day services'activities for persons with profound intellectual and multiple disabilities. J Policy Practice Int Dis. 2007;4:152-159.
29 Luckasson R, Borthwick-Duffy S, Buntinx W, Coulter D, Craig P, Reeve A. et al. Mental Retardation:Definition, Classification and Systems of Supports. 2002, American Association on Mental Retardation, Washington.
30 Schalock R, Brown I, Brown R, Cummins RA, Felce D, Matikka L, Keith KD, Parmenter T. Conceptualization, Measurement, and Application of Quality of Life for Persons With Intellectual Disabilities: Report of an International Panel of Experts. Ment Retard 2002;40(6):457-470.
31 Dutch Society for Doctors in Revalidation, 2006, Protocol for diagnosis and treatment of children with cerebral palsy and spasticity.
32 Dutch Society for Doctors in the Care for people with an Intellectual Disability (NVAZ) Resistance among people with an intellectual disability in the framework of the Act Governing Medical-Scientific Research Involving Humans, behavioural code for doctors in the assessment of resistance among people with an intellectual disability, 1999.
33 Karvonen MJ, Kentala E, Mustala, O. The effects of training on heart rate: a longitudinal study. Ann Med Exp Fenn. 1957;35(3):307-315.
34 Fernhall B, Millar AL, Tymeson GT, Burkett LN. Maximal exercise testing of mentally retarded adolescents and adults: reliability study. Arch Phys Med Rehabil. 1990;71(13):1065-8.
35 Brennan RL. Generalizability theory. 2001, New York: Springer-Verlag
36 Vermunt JK & Magidsen J. Latent class models for classification. Computational Statistics & Data Analysis. 2003;41(3-4):531-537
37 Vermunt J., Tran B, Magidsen J. Latent class models in longitudinal research. In: Menard S, editor. Handbook of longitudinal research: design, measurement, and analysis Burlington: Elsevier. 2007:373-385.

38 Multilevel Models Project (2004) A user's guide to MLwiN [computer program]. Version 2.10. London.
39 Putten AAJ van der & C Vlaskamp. Motor activation in people with PIMD. J App Res Int Dis. 2010;23:496.
40 Vlaskamp C & Putten AAJ van der. Physical activities in people with PIMD: when and how are they offered. J App Res Int Dis. 2010;23: 496.
41 Lemmer B. Importance of circadian rhythms for regulation of the cardiovascular system-studies in animal and man. IEEE Engineering in Medicine and Biology, Annual Conference 1. 2006;168-70.
42 Schroeder A, Loh DH, Jordan MC, Roos KP, Colwell CS. Circadian Regulation of Cardiovascular Function: a role for vasoactive intestinal peptide. J Am Phys Heart Circ Phys. 2010; 15. [Epub ahead of print]
43 Sijtsema JJ, Veenstra R, Lindenberg S, Van Roon AM, Verhulst FC, Ormel J, Riese H. Mediation of Sensation Seeking and Behavioral Inhibition on the Relationship Between Heart Rate and Antisocial Behavior: The TRAILS Study. J Am Acad Child Adolesc Psychiatry. 2010;49:493-501.
44 Ortiz J & Raine A. Heart rate level and antisocial behavior in children and adolescents: a meta-analysis. J Am Acad Child Adolesc Psychiatry. 2004;43:154-162.
45 Raine A. Annotation: the role of prefrontal deficits, low autonomic arousal, and early health factors in the development of antisocial and aggressive behavior in children. J Child Psychol Psychiatry. 2002;43:417-434.
46 Vos P, De Cock P, Petry K, Van Den Noortgate W, Maes B. Do You Know What I Feel? A First Step Towards a Physiological Measure of the Subjective Well-Being of Persons With Profound Intellectual and Multiple Disabilities. J App Res Int Dis. 2010;23:366–378.

Chapter 9

General Discussion

As this thesis is comprised of multiple studies related to psychometric characteristics of measurements concepts when testing for a multiple disabled target group, a solid overview of the results and their relations is a necessary and natural last step. This discussion is roughly divided into four parts. The first part touches upon general findings, which deal with classification, test setting and adjustments in formulas. Then, this discussion goes into more detail discussing results grouped per studied concept: body composition, functional exercise and aerobic capacity, balance, flexibility and physical activity respectively. Additionally, the third part deals with the methodological issues encountered during research; lack of standards and cutt-off scores, small sample size and heterogeneity of the study population. The final part of this discussion seeks to give an overview of the implications of the different studies comprised in this thesis, as well as to indicate recommended areas of further research.

General Findings

Classification and Test Setting

The subjects of this thesis have severe or profound intellectual as well as visual disabilities and the study population is thus referred to as persons with severe or profound intellectual and multiple disabilities (SPIMD). The term 'multiple' includes locomotor disabilities, neurological problems, sensory disabilities, and/or problems with food ingestion.

An important topic that came up during the studies of this thesis [all chapters] is the importance of combining both severity of intellectual disabilities as locomotor levels (GMFCS) to classify the abilities of persons with severe or profound intellectual disabilities. Only if both classifications are used, the appropriate physical fitness tests can be individually selected. Contrary to common assumptions, persons with profound intellectual disabilities do not automatically have low locomotor levels. The ability to walk varies considerably in persons with severe intellectual disabilities as well as in those with profound intellectual disabilities. Specifically, 75 % of the subjects with severe intellectual disabilities were able to walk, leaving 25 % which were not able to walk. In contrast, 56 % of the subjects with profound intellectual disabilities is able to walk, whereas 44 % is not able to walk [1, 2]. Not testing for GMFCS levels thus means ignoring a big potential for both research as improvement of individiual physical fitness.

Also, contrary to common beliefs, persons with SPIMD can become accustomed to test and measurement situations despite their severe or profound intellectual and multiple disabilities, if optimal test environment and conditions are created. Testing in persons with SPIMD in a feasible, valid and reliable way is possible –albeit challenging due to the limitations in intellectual functioning, adaptive behaviour [3] visual impairment [4], low motivation and adherent stress [5, 6].

This thesis made use of environmental cues to counter these limitations. Environmental cues included adaptations of existing test protocols, performing practice sessions with a familiarization protocol, testing at the regular physical activity hours in the regular sports hall accompanied by familiar gymnastic instructors, who were used to motivate the participants. In addition, the use of a treadmill to assess cardiorespiratory fitness is shown to be an effective environmental cue. Following theories about perception-action coupling [7, 8], these environmental cues may have facilitated the performance of the subjects by both reducing stress as enhancing motivation. Technology could be used to develop other environmental cues such as auditive or other sensory stimuli, which may even further ameliorate the testing situation. Future research should focus

on this. To sum up, testing in persons with SPIMD in a feasible, valid and reliable way is possible, if participants are classified according to both GMFCS levels and level of disability, and if optimal test environment and conditions are created.

Adjusted Formulas

This thesis has brought forth a couple of important findings regarding formulas or equations applicable to research on this specific target group. First of all, it proposes a new reliable equation to calculate body height from tibia length [1]. Second, a simple prediction equation is suggested for predicting standing waist circumference from supine waist circumference measurement [2]. The latter equation allows the comparison of supine measurements of waist circumference in disabled persons with international standards [9]. Third, we also found that the Fernhall's formula [10, 11] for calculating peak heart rate for people with ID, systematically overestimates peak heart rate for people with SIMD [chapter 5]. The subheader "Functional exercise and aerobic capacity" discusses this finding in more detail.

Overall, whereas for other specific target groups such as the elderly, children with cerebral palsy, and persons with mild or moderate ID feasible and reliable fitness tests already were available [5, 12, 13, 14, 15], the feasibility and reliability of physical fitness measurements and tests for participants with SPIMD had so far not been established. Using the feasible and reliable fitness tests described in this thesis, physical fitness levels can now be objectively evaluated. Moreover, specific interventions aimed at promoting physical fitness of participants with SPIMD can be evaluated objectively as well.

Not only does this study thus contribute to existing literature, it also, quite practically delivers tools for the work floor. For instance, the fitness tests that were evaluated in chapters 2 to 8 of this thesis have proven to be feasible, valid and reliable for persons with severe or profound intellectual and multiple disabilities, and could thus immediately be implemented in daily practice.

What follows is a more detailed overview of the results per studied fitness concept.

Body composition

Anthropometric measurements are widely used to reliably quantify body composition and to estimate risk of overweight in both healthy subjects and patients. Body composition of persons with ID is described in several studies [16, 17, 18]. However, information about the reliability of anthropometric measurements in persons with SIMD had so far not been subject to research. In 45 participants with SIMD, body mass index, waist circumference, skinfolds, and tibia length were measured twice [1]. The results of this study show that for individuals with SIMD and GMFCS levels I and II, feasibility and test-retest reliability for body composition measurements were acceptable, with exception of the skinfold measurements. Body Mass Index is calculated using body height and body weight. However, body height is difficult to measure for those subjects unable to stand upright. Therefore, a study was set up in order to calculate body height from tibia length. This study [1, chapter 2] revealed a new reliable equation to calculate body height from tibia length. Measuring tibia length is more feasible, accurate, and reliable than measuring the total body height of participants with SIMD- both for those able to stand as for those not able to do so.

Waist circumference as an indicator of abdominal fat is an important predictor of health risks and is usually measured in standing position. We found that measuring waist circumference in participants with SIMD who are able to stand upright is feasible and reliable. However,

due to motor disabilities, many individuals with PIMD and GMFCS levels IV and V are unable to stand straight or to stand at all. When dealing with these participants, waist circumference can only be measured with the participant lying in a supine position. It had so far been unknown whether measuring waist circumference of a participant in a supine position is valid and reliable. This issue is particularly relevant when international standards [10] for healthy individuals are applied to the disabled. The results of a test-retest study [2, chapter 3] performed in 43 participants with PIMD and GMFCS levels IV and V indicated that supine waist circumference can be reliably measured for this target group. Our validity study [2, chapter 3] performed in 160 healthy participants, compared waist circumference obtained in both standing and supine positions. This study shows that supine waist circumference is biased towards higher values (1.5 cm) when compared with standing waist circumference. A simple equation could be put forward, which predicts standing waist circumference based on supine measurements. Such an equation allows for the comparison of waist circumference measurements of disabled persons with the international standards [10]. Based on BMI and waist circumference values, 10% of the female SPIMD subjects are obese and 39% are abdominal obese, while 0% of the male persons are obese and only 7% are abdominal obese. Thus, women with SPIMD are at a higher risk for developing overweight related health problems compared to SPIMD classified men. Compared to outcomes of measurements in persons with mild or moderate ID, persons with SPIMD show a more healthy body composition status [18, 19, 20]. Further research into the explaining mechanisms behind this interesting observation is recommended.

Another issue related to body composition stems from the fact that children and adults who have SPIMD and GMFCS level IV or V are often fed by stomach tube [21]. Tube feeding may increase body weight mainly through fat deposition [22]. Children and adults with severe CP have relatively high body-fat/muscle-content ratio. When the relatively low energy expenditure is also taken into account, we see a potential risk of overfeeding [23]. To establish a healthy body weight in persons with PIMD, it is necessary to take into account both BMI and waist circumference. Rieken et al [24] suggest to use bioelectrical impedance analysis as an alternative measurement of body composition and present a preliminary clinical guideline on diagnosing undernutrition and overnutrition in children with severe neurological impairment and ID. We recommend that a similar guideline will be designed for adults with SPIMD and GMFCS levels IV and V, so as to prevent over- or undernutrution in this target group.

Functional exercise and aerobic capacity

Cardiorespiratory fitness can be divided into functional exercise and aerobic capacity [35]. Several tests are available for measuring functional exercise and aerobic capacity, including the six-minute walking test (6 MWD) and the shuttle run test (SRT). However, it was unclear whether these tests are feasible and reliable when testing subjects with SIMD.

Our studies [10, chapter 4] showed that the 6 MWD is feasible and reliable for testing participants with SIMD and GMFCS levels I and II. An adapted SRT (aSRT) performed overground as well as on a treadmill has proven to be feasible and reliable for a target group with SIMD and GMFCS level I [chapters 4 and 5]. Moreover, the aSRT performed on a treadmill [chapter 5] appeared to be more valid for determining peak heart rate than the aSRT performed overground for people with SIMD and GMFCS level I. We also found that the Fernhall's formula [11] for calculating peak heart rate of subjects with ID systematically overestimates peak heart rate of subjects classified with SIMD. Thus,

for future research it is recommended to adjust this equation so as to ensure a valid prediction of the maximal heart of this specific group.

Furthermore, we compared the achieved mean distance in the 6MWD of our participants with values reported in other studies [25, 26, 27]. This comparison indicated that persons with SIMD perform poorer on the 6MWD than those with other specific (chronic) health conditions including heart failure or COPD. The poor 6MWD results we observed suggest that the low functional exercise capacity of individuals with SIMD is a serious health problem, which in turn can negatively affect their independence in day-to-day activities.

Based on this result, further research should be directed towards developing, implementing and evaluating interventions aimed at increasing functional exercise and aerobic capacity of SIMD classified subjects thereby reducing related health problems.

Balance

Sufficient balance is necessary to perform daily activities. However, individuals with SIMD are particularly at risk when it comes to the development of deficits in locomotor skills [28] and are likely to have decreased balance. Yet, the feasibility and reliability of balance tests for subjects with SIMD had so far not been subject of research. The Berg Balance Scale (BSS) is a widely used test to quantify balance. Our study [29, chapter 6] revealed a sufficient percentage of successful measurements by the modified BBS (mBBS), indicating this test to be a feasible instrument for testing subjects with SIMD and GMFCS levels I and II. In addition, with exception of two out of the twelve mBBS tasks to be performed by the participants, reliability for test and retest appeared to be sufficient.

In the BBS, a score of 80% (45 points) indicates sufficient balance [30]. As its modifications influence the cut-off value, it was impossible to apply the same cut-off value to the mBBS. A comparison with scores in other target groups was hence not possible. Future research should be directed towards establishing accurate cut-off values.

Flexibility

Muscle flexibility is a relevant part of physical fitness and can be measured using the Modified Ashworth Scale (MAS) and the Modified Tardieu Scale (MTS). The results of our study [31, chapter 7] showed the feasibility of the MAS and MTS to be sufficient for testing participants with PIMD. For both inter and intrarater reliability, measurements of the MAS revealed acceptable agreement. However, the measurements of the MTS showed insufficient agreement for both test-retest and interrater reliability.

For our target group the MAS showed a better test-retest and interrater reliability than the MTS, which contradicts the results of both Gielen [32] and Mehrholz et al. [33] who found the opposite to be true for subjects with respectively ID and adult patients with severe brain injury. In their studies the psychometric properties of the MTS were better than those of the MAS. Our measurements of the MTS however, put forward LOAs not narrow enough to indicate agreement between the two measurements of the same subject, leading us to conclude the LOAs for the MTS measurements to be clinically unacceptable.

Other authors [34] have suggested the MTS to be a more appropriate clinical measure of spasticity than the MAS. We thus assessed whether the LOA of the MTS measurements of our subjects with spasticity was narrower than that of our subjects without spasticity.

Yet, no difference was found between these groups. For our target group, including those suffering from spasticity, MAS seems the preferred measurement instrument. This preference may be accounted for by the interaction beween multiple disabilities in our study population, which may differ in that respect from the study populations those of Mehrholz [34] and Gielen [33].

However, the validity of the aforementioned instruments for measuring spasticity and muscle tone is beyond the scope of this study. We recommend the validity of the MAS to be examined in future studies, with a focus on measuring the quality of daily movement rather than measuring spasticity, as the MAS may be more suitable for measuring the first.

Physical activity

Daily physical activity contributes directly to health [35] and is an important factor in maintaining and improving physical fitness. Monitoring daily physical fitness in persons with SPIMD is thus an important step in identifying those at health risks.

Our research [chapter 8] has shown that daily physical activity can be reliably monitored through measuring heart rate patterns. Yet, further exploration of other influential factors such as emotions [36] is still recommended. Time of day and age have considerable influence on heart rate patterns. However, the analysis of heart rate patterns suggests other, probably more personal factors to have a significant influence on heart rate patterns [37] which can be explored in future research. Finally, our study shows that based on measuring heart rate patterns, individuals with PIMD do not meet the standards of sufficient activity as proposed by the guidelines of the American College of Sports Medicine [38]. Future research can shed a light on the suitability and validity of these guidelines for a population with PIMD or SPIMD.

Methodological issues

The first methodological issue we encountered had to do with the lack of a golden standard. As validity can only be examined against a golden standard, it would have been convenient and desirable for a golden standard to be available for every single instrument measuring a certain concept. Yet realistically, golden standards for our research population are unfortunately simply not available. If feasible, we attempted to assess validity, for example by performing a supra maximal block test on the treadmill.
Moreover, often neither standard values nor cut-off scores are available, as was the case for the aSRT nor the mBBS.

Another methodological issue is related to statistical power. Due to both the relatively small group of persons with SPIMD and a number of exclusion criteria, a rather small number of subjects were able to participate in the studies described in this thesis. However, despite the relatively small study groups the test-retest reliability of mBBS, walking tests, bodycomposition measurements and the MAS were comparable to those of other and larger target groups. In future research, more care facilities and thus subjects should be included and classified on both of level of ID and locomotor skills. This may improve sample size and thereby power. Then it might be possible to formulate golden standards, group specific standard values and cut-off scores, with which the validity of tests can be examined and the testscores can be compared.

Despite the fact that only cross-sectional designs are used in this research, we will touch upon the problems of heterogeniety in samples with SPIMD in the framework of intervention

studies. Randomized Controlled Trials (RCT's) are considered to be the most reliable and valid way to perform intervention studies in various populations. Comparison between groups is thought to reflect differences in effect of the interventions. The required sample size in RCT's depends partly on the variation between participants: a large variation calls for a larger sample size. Therefore, many studies include homogeneous groups. However, due to the variety of their co-morbidities, it is difficult to compose a homogeneous study population with sufficient power composed of persons with SPIMD. Neither is it possible to compose a large heterogeneous SPIMD study population. As a consequence, both the experimental and control group will consist of participants with much variation in their co-morbidities, resulting in a wide array of responses to the same intervention. In such circumstances detecting significant effects of an intervention is rather difficult. An anecdote from practical experience will illustrate this.

To subject three participants with SPIMD to passive or assistive active movements, powered exercise equipment (Shapemaster®, Barth Fidder, Shapemaster Benelux) was used. These machines are fitted with motors and gearboxes, and controlled by microchip technology. The machines automatically move selected levers and handles at pre-determined speeds through a pre-determined range of motion. Each machine provides multi-function movements. The outcome measures are bodycomposition, muscle tone, heart rate, oxygen saturation and alertness [39]. At the individual level, relevant improvements were found for the different outcome measures.

1. A woman of 38 years old, with profound ID, GMFCS level IV, no spasticity, totally blind, epilepsy, and being overweight, participated in the study. Her BMI before the intervention was 27.7 kg/cm^2, after the intervention 26.2 kg/cm^2, which means a difference of 1.5 kg/cm^2. Her waist circumference decreases from 89 cm, which means abdominal obesitas, to 83 cm, which is indicating 'healthy weight but attention needed' [7]. Oxygen saturation during and after moving on the machines increased from 95% before the intervention program, to 99 % after 20 weeks. However, muscle tone, alertness and hart frequency showed no differences.
2. A girl of 17 years old, with profound ID, GMFCS level V, severe partially sighted, with spasticity, epilepsy and orthopedic defects also participated in the study. Her muscle tone in the legs decreased in 20 weeks with two points on the six point scale of the Modified Ashworth Scale. After every increase in intensity in a five weeks period, her heart rate increased first one or two heart rate zones during moving, but after three weeks, heart rate decreased again to the first level. This might indicate a training effect. However, saturation, alertness, BMI and waist circumference showed no differences.
3. A man of 43 years old, with profound ID, GMFCS level V, totally blind, spasticity, epilepsy, and orthopedic defects, also participated in the study. His muscle tone in both arms and legs decreased in 20 weeks with one point on the six point scale of the Modified Ashworth Scale. Oxygen saturation during moving on the machines increased from 91% before the intervention program, to 95 % after 20 weeks. BMI decreased after the intervention with 0.5 kg/cm^2, but before and after the intervention he already had a healthy BMI. Alertness increased during intervention with one point on a four point scale [39]. However, heart rate showed no differences during and after the intervention period.

As shown, individuals benefit from the intervention but not in the same way nor to the same extent. Individual differences in characteristics of locomotor skills, visual impairment, co-morbidities, and baseline measurements of the outcome measures account for this result. Yet, there were benefits, albeit different ones for different subjects. In group comparison studies the relevant individual benefits can be overlooked. Consequently, next to traditional research designs,

alternative research designs such as multiple case studies or program evaluation and adherent statistical analysis [40, 41] should be used for intervention studies aimed at participants with SPIMD.

Implications

In the introduction we presented an integration of models and concepts describing quality of life, participation, physical well-being, physical fitness, physical activity and health, so as to illustrate their mutual relatedness (figure 1).

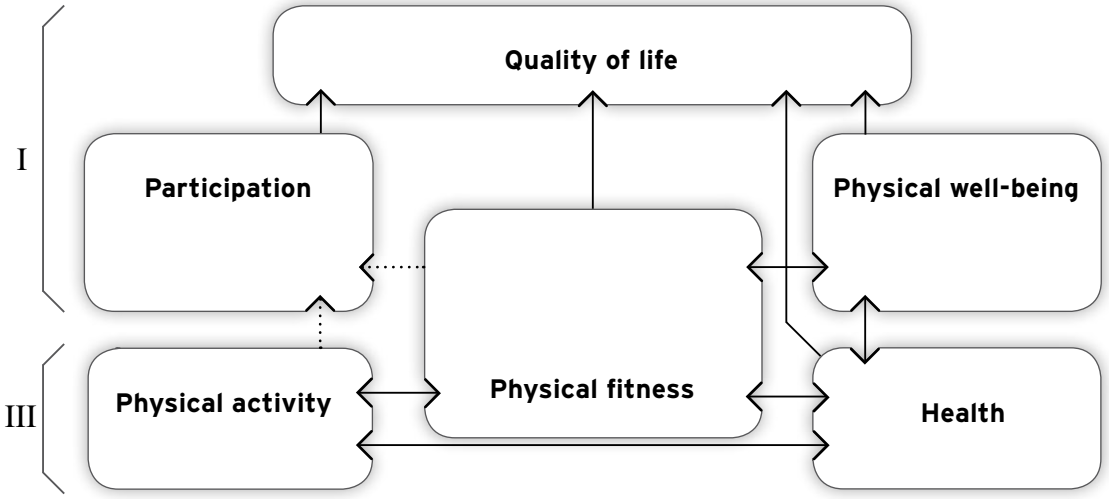


Figure 1. Integration of models and concepts of participation, quality of life, physical well-being, physical activity, physical fitness, and health [3, 35, 42, 43, 44].

As we put forward in the introduction, out of the concepts introduced in this model, this thesis mainly focused on the measurement of physical activity and fitness, operationalized by body composition, functional exercise and aerobic capacity, balance, muscle flexibility and heart rate patterns.

Now that physical fitness and physical activity can be evaluated in persons with SPIMD, the next step to take in future research is to look at the relation of physical activity and fitness with the concepts on the right hand side of the figure, including health, physical well-being, and quality of life. Also, further research should elaborate on the relation between phyiscal fitness and the concepts in the left hand side of the figure, including participation. This however can be a hard nut to crack, as evidence on the concept of participation is lacking. The direct relation between physical activity / fitness and participation in persons with SPIMD has so far been unknown as the level of participation is difficult to define and quantify for this specific group[45]. The question is whether persons with SPIMD are really able to participate in daily activities. Can they take part in or have influence on situations and contexts important to them personally? A complicating factor in this respect is the fact that these persons are often not able to verbalize what is important to them and, what's more, to verbalize what to them 'taking part' or 'having influence' actually means. It is often observed that a person merely sitting in a wheelchair placed in an activity room with others, is considered to 'take part' in that activity or society [45]. Or does ''participation''

of these individuals require more? It is imperative and essential to investigate to which extent a person with SPIMD really participates in living habits, 'work', leisure activities, sports etc. However, so far knowledge on this topic is not available neither are instruments aimed at investigating the concept of participation in persons with SPIMD.

Recommendations

The vast majority of recommendations for future research have been mentioned previously- either under the studied concepts or in general findings. These however all amount to a call for tailored interventions. Interventions aiming at promoting physical fitness are generally carried out with participants suffering from mild or moderate ID. Performing similar interventions with participants classified as having severe or profound ID is thought to be more difficult, firstly because of the assumption that most of these persons are not able to walk, and secondly because these persons have more problems understanding the tasks required of them. We recommend developing, performing and evaluating tailored interventions geared to promote physical activity and fitness in persons with severe or profound ID who may or may not be able to walk. As environmental cues facilitate performance in persons with SPIMD, examining environmental cues provided by technological adaptations, like auditive or other pleasant sensory stimuli can be of significant help.

As knowledge on the concept of participation of individuals with SPIMD is lacking, further research should aim on feasible instruments to identify and quantify outcome measures of participation. Furthermore, research aimed on exploring the relations between physical fitness and physical activity on one side, and participation on the other, is recommended.

Previous paragraphs have discussed the shortcomings, questions and suggestions that came up during the various studies comprised by this thesis. This thesis functions as a basic first step to enhance and strenghten the role of physical activity and fitness for this target group, by enabling sound academic testing of physical fitness. We strongly advocate to consider the major role physical activity and fitness can play in well-being.

References

1 Waninge A, van der Weide W, Evenhuis IJ, van Wijck R, van der Schans CP. Feasibility and reliability of body composition measurements in adults with severe intellectual and sensory disabilities. J Intellect Disabil Res. 2009;53(4):377-88.

2 Waninge A, Ligthart KA, Kramer J, Hoeve S, van der Schans CP, Haisma HH. Measuring waist circumference in disabled adults. Res in Dev Dis. 2010;31(3):839-47.

3 Schalock R, Brown I, Brown R, Cummins RA, Felce D, Matikka L, Keith KD, Parmenter T. Conceptualization, Measurement, and Application of Quality of Life for Persons With Intellectual Disabilities: Report of an International Panel of Experts. Ment Retard. 2002;40(6):457-470.

4 Van Splunder J, Stilma JS, Bernsen RM & Evenhuis HM. Prevalence of visual impairment in adults with intellectual disabilities in the Netherlands: cross-sectional study. Eye (Lond). 2006;9:1004-10.

5 Rintala P, Dunn JM, McCubbin JA, Quinn C. Validity of a cardiorespiratory fitness test for men with mental retardation. Med Sci Sports Exerc 1992;24(8):941-5.

6 Frey GC, Buchanan AM, Rosser Sandt DD. "I'd rather watch TV": an examination of physical activity in adults with mental retardation. Ment Retard. 2005;43(4):241-54.

7 Newell KM. Constraints on the development of coordination. In Motor Development in Children: Aspects of Coordination and Control . In: M.G.Wade & H.T.A.Whiting, editor. Dordrecht, Netherlands: Martinus Nijhoff;.1986;341-60.

8 Davids K, Button C, Bennett S. Dynamics of Skill Acquisition, a constraints-led approach. 2008, Human Kinetics, Champaign, IL.

9 World Health Organization, Waist circumference. 2006, World Health Organization.

10 Waninge A., Evenhuis IJ, van Wijck R, van der Schans CP. Feasibility and reliability of two different walking tests in subjects with severe intellectual and sensory disabilities. Accepted for publication in Journal of Applied Research in Intellectual Disabilities, July 2010

11 Fernhall B, McCubbin J A, Pitetti KH, Rintala P, Rimmer JH, Millar AL, De Silva A. Prediction of maximal heart rate in individuals with mental retardation. Med Sci Sports Exerc. 2001;33:1655-1660.

12 Millar AL, Fernhall B, Burkett LN. Effects of aerobic training in adolescents with Down syndrome. Med Sci Sports Exerc.1993;25(2):270-4.

13 Verschuren O, Takken T, Ketelaar M, Gorter JW, Helders PJM. Reliability and validity of data for 2 newly developed shuttle run tests in children with cerebral palsy. Physical Therapy. 2006;86:1107-1117.

14 Troosters T, Gosselink R, Decramer M. Six minute walking distance in healthy elderly subjects. Eur Resp Journal. 1999;14:270-274.

15 Berg K. Balance and its measure in the elderly: a review. Phys Can. 1989;41:240-6.

16 Bhaumik S, Watson JM, Thorp CF, Tyrer F, Mc Grother CW. Body mass index in adults with intellectual disability: distribution, association and service implications: a population-based prevalence study. J Intellect Disabil Res. 2008;52(Pt 4):287-98.

17 Frey GC & Chow B. Relationship between BMI, physical fitness, and motor skills in youth with mild intellectual disabilities. Int J Obes. 2006;30:861-867.

18 Knijff-Raeven van AGM, Jansen-Jacobs CCM, Freen PJW, Hoekman J, Maaskant MA. Body Mass Index (BMI) bij mensen met een verstandelijke beperking. Nederlands Tijdschrift voor de Zorg aan mensen met verstandelijke beperkingen. 2005;1:3-17.

19 Lahtinen U, Rintala P, Malin A. Physical performance of individuals with intellectual disability: a 30 year follow up. Adapt Phys Act Q. 2007;24: 125-143.
20 McGuire BE, Daly P, Smyth F. Lifestyle and health behaviours of adults with an intellectual disability. J Intellect Disabil Res. 2007;51(Pt 7):497-510.
21 Stevenson RD, Conaway M, Chumlea WC, Rosenbaum, O'Donnell, Samson-Fang L, Stallings VA, Fung EB, Henderson RC, Worley G, Liptak G. Growth and Health in Children With Moderate-to-Severe Cerebral Palsy. Pediatrics. 2006;118:1010-1018
22 Kong CK & Wong HS. Weight-for-height values and limb anthropometric composition of tube-fed children with quadriplegic cerebral palsy. Pediatrics. 2005;116, 839-845.
23 Sullivan PB, Alder N, Bachlet AM Grant H, Juszczak E, Henry J, Vernon-Roberts A, Warner J, Wells J. Gastrostomy feeding in cerebral palsy: Too much of a good thing. Dev Med Child Neurol. 2006;48:877-82.
24 Rieken R. Assessing body composition and energy expenditure in children with severe neurological impairment and intellectual disability. 2010, Optima Grafische Communicatie, Rotterdam.
25 Troosters T, Vilaro J, Rabinovich R, Casas A, Barbera JA, Rodriguez-Roisin R. Physiological responses to the 6-min walk test in patient with chronic obstructive pulmonary disease. Eur Resp Journal. 2002;20:564-569.
26 Faggiano P, D'Aloia A, Gualena A, Lavatelli A, Giordano A. Assessment of oxygen uptake during the 6-minute walking test in patients with heart failure: preliminary experience with a portable device. Am Heart Journal.1997;134: 203-206.
27 Onorati P, Antonucci R, Valli G, Berton E, De Marco F, Serra P. Non-invasive evaluation of gas exchange during a shuttle walking test vs. a 6-min walking test to assess exercise tolerance in COPD patients. Eur J Appl Phys. 2003;89:331-336.
28 Evenhuis HM, Sjoukes L, Koot HM, Kooijman AC. Does visual impairment lead to additional disability in adults with intellectual disabilities? J Intellect Disabil Res. 2009;53(1):19-28.
29 Waninge A, van Wijck R, Steenbergen B, van der Schans CP. Feasibility and reliability of a modified Berg Balance Scale in persons with severe intellectual and sensory disabilities. J.Int Disabil Res. 2011;55(3):292-301.
30 Berg K, Wood-Dauphinee S, Williams JI, Maki B. Measuring Balance in the Elderly: Preliminary development of an Instrument. Physiotherapy Canada. 1989;41:304-311.
31 Waninge A, Rook RA, Dijkhuizen A, Gielen E, van der Schans CP. Feasibility, test-retest reliability and interrater reliability of the Modified Ashworth Scale and Modified Tardieu Scale in persons with profound intellectual and multiple disabilities. Res Dev Dis. 2011;32(2):613-620.
32 Gielen EJJM. Is spasticiteit te meten bij mensen met een verstandelijke beperking? Een betrouwbaarheids- en validiteitsonderzoek van de Modified Ashworth Scale en de Tardieu Schaal bij mensen met een verstandelijke beperking. Scriptie in het kader van de cursus Scholing in Wetenschap III; Amersfoort, Nederlands Paramedisch Instituut, 2005.
33 Mehrholz J, Wagner K, Meiβner D, Grundmann K, Zange C, Koch R, Pohl M. Reliability of the Modified Tardieu Scale and the Modified Ashworth Scale in adults patients with severe brain injury: a comparison study. Clinic Rehabili. 2005;19:751.
34 Haugh AB, Pandyan AD, Johnson GR. A systematic review of the Tardieu Scale for the measurement of spasticity. Dis Rehabil. 2006;28:899-907.

35 Bouchard C, Shepard RJ, Stephens T. Physical activity, Fitness and Health. 1994, Human Kinetics Publishers, ChampaignIL.
36 Vos P, De Cock P, Petry K, Van Den Noortgate W, Maes B. Do You Know What I Feel? A First Step Towards a Physiological Measure of the Subjective Well-Being of Persons With Profound Intellectual and Multiple Disabilities. J App Res Int Dis. 2010;23:366–378.
37 Raine A. Annotation: the role of prefrontal deficits, low autonomic arousal, and early health factors in the development of antisocial and aggressive behavior in children. J Child Psychol Psychiatry. 2002;43:417-434.
38 American College of Sports Medicine Position Stand. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. Med Sci Sports Exerc. 1998;30: 975-91.
39 Munde VS, Vlaskamp C, Ruijssenaars AJ, Nakken H. Alertness in individuals with profound intellectual and multiple disabilities: a literature review. Res Dev Disabil. 2009;30(3):462-80.
40 St Pierre RG, Rossi PH. Randomize groups, not individuals: a strategy for improving early childhood programs. Eval Rev. 2006;30(5):656-85.
41 Yin RK. Enhancing the quality of case studies in health services research. Health Serv Res. 1999;34(5 Pt 2):1209-24.
42 Kramer AF, Hahn S, Cohen NJ, Banich MT, McAyley E, Harrison CR, Chason J, Vakil E, Bardell L, Boileau RA, Colcombe A. Ageing, fitness and neurocognitive function. Nature. 1999;400:418-419
43 Sitskoorn, MM. Het plastische brein; invloed van gedrag. De Psycholoog. 2005;40(5): 262-267.
44 Colcombe S, Erickson KI, Raz N, Webb AG, MC Auley EB, Kramer AF. Aerobic fitness reduces brain tissue loss in aging humans. J Gerontol. 2003;58:176-180.
45 Kiestra T. (2005), De unieke handicap, referentiemodel voor meervoudige beperkingen, Bedum, Scholma Druk, ISBN 9080638420.

Summary

Introduction

Physical fitness, physical activity and health are related in the sense that good physical fitness and a sufficient level of physical activity may improve both mental and physical health. In turn, good physical fitness and health may have a positive effect on wellbeing, participation, and quality of life.

Therefore, it is imperative to gain a comprehensive insight into physical fitness and physical activity levels of persons with severe or profound intellectual and visual (multiple) disabilities (SPIMD). These persons are at risk to develop health problems and are in danger of being excluded from many situations and opportunities usually available to people not suffering from intellectual disabilities (ID). However, due to limitations in intellectual functioning and adaptive behaviour on one hand, and visual and locomotor functioning on the other hand, the level of health-related physical fitness of persons with SPIMD is difficult to quantify in a feasible, valid and reliable manner. As a consequence, knowledge of the physical fitness levels and locomotor skills of persons with SPIMD is scarce.

Thus, the purpose of this thesis is to examine the feasibility, validity and reliability of physical fitness tests for persons with SPIMD.

Persons with intellectual disabilities

Four levels of ID can be distinguished: mild, moderate, severe or profound. Adults with severe ID have an intellectual age ranging from 3 to 6 years, which is likely to result in a continuous need for support. Adults with profound ID have an intellectual age below 3 years, which results in serious limitations in self-care, continence, communication and mobility.

When compared to the general population, persons with ID have twice as much health problems and significantly higher levels of co-morbidity such as epilepsy, neurological problems, visual or auditive impairment and locomotor disabilities. Moreover, the level of co-morbidity in persons with severe or profound ID is even higher than co-morbidity levels in persons with mild or moderate ID.

Research showed that persons with ID display inadequacies in perception, motor-reproduction and sensorimotor control. Furthermore, persons with ID are often not sufficiently active to achieve health benefits, and more than 50 % of the persons with ID of all age categories in Europe have a sedentary lifestyle. In addition, those classified with ID are more prone to experience lifestyle related diseases such as diabetes mellitus II or cardiovascular diseases. These persons often suffer from overweight or malnutrition. As a consequence, these persons may have poor physical fitness.

Physical fitness of persons with severe or profound intellectual and visual disabilities

Similar to individuals with ID, persons with visual impairments display poor performance on locomotor skills and have low levels of habitual activity, resulting in poor physical fitness when compared to the control group, in this case persons with normal eyesight.

Prevalence of visual impairment in persons with severe or profound ID is 92%. As the combination of ID and visual impairment is even more detrimental, thereby creating less opportunity for compensation, the combination of visual impairment with ID aggravates

problems in both locomotor skills as in daily functioning. These findings together put forward the suggestion that persons having SPIMD are likely to display insufficient physical fitness.

However, the feasibility and reliability of physical fitness measurements and tests for participants with SPIMD have until now not been properly scrutinized, resulting in little reliable knowledge on the physical fitness levels and locomotor skills of persons with SPIMD.
Due to the limitations related to SPIMD, the level of health-related physical fitness is difficult to quantify in a feasible, valid and reliable manner. Therefore, improving feasibility of physical fitness tests for individuals with SPIMD needs to be prioritized. Persons with SPIMD are not accustomed to assessments and have difficulty comprehending what is required of them. Furthermore, persons with visual disabilities cannot see how test tasks need to be performed, hence showing them how to perform the task at hand is useless. In general, if a participant does not understand the tasks within a certain test, the test will automatically fail to provide a realistic impression of the capabilities of the participant, rendering the test invalid. Thus, test instructions for persons with SPIMD require our special focus.

Another factor of influence when determining the feasibility, reliability and validity of physical fitness tests for participants with SPIMD is the prevalence of locomotor disabilities and motivational problems. As persons with SPIMD are not able to stand straight or to stand at all, adapted test procedures and specific inclusion criteria are required. Also, persons with SPIMD are often not motivated to exert themselves fully, which necessitates adjustments to and familiarization with test protocols.

The required attributes of physical fitness for persons with SPIMD described by caregivers, professionals and scientists in the field of SPIMD are: body composition, cardiorespiratory fitness, balance, muscle strength and muscle flexibility.

As stated before, the locomotor skills of persons with ID vary considerably and this may influence protocols for measuring physical fitness. Thus, next to the level of ID, persons with SPIMD are also grouped according to a motor function classification used to classify the locomotor skills in people with physical disabilities. In the studies examining a population consisting in majority of persons with severe intellectual disabilities, the term severe intellectual and multiple disabilities (SIMD) is used. In the studies examining a population consisting in majority of persons with profound intellectual disabilities, the term profound intellectual and multiple disabilities (PIMD) is used.

Aim and research questions of the thesis

The main aim of the research reported in this thesis is to examine the feasibility, validity and reliability of physical fitness tests for individuals with SPIMD.

This research addresses the following research questions:

1. Are body composition measurements of participants with SIMD feasible and reliable [chapter 2]?
2. Are waist circumference measurements of participants with PIMD valid and reliable [chapter 3]?
3. Is cardiorespiratory fitness of persons with SIMD measured feasibly and reliably with either or both of two different walking tests? [chapter 4]?
4. Is a walking test performed on a treadmill feasible, valid and reliable for persons with SIMD [chapter 5]?
5. Is a balancetest feasible and reliable for persons with SIMD [chapter 6]?
6. Are two tests to measure muscle tone and spasticity feasible and reliable when testing persons with PIMD [chapter 7]?
7. What is the level of physical activity of persons with PIMD based on heart rate patterns when compared to the American College of Sports Medicine guidelines of healthy physical activity? Is there a relation between heart rate patterns and observed level of activity in persons with PIMD? What is the influence of covariates such as gender, age, and common co-morbidity (motor disabilities, spasticity and sensory disabilities) on heart rate patterns [chapter 8]?

Body composition measurements

Anthropometric measurements are widely used to reliably quantify body composition and to estimate risks of overweight in both healthy subjects as in patients. However, information about the reliability of anthropometric measurements of persons with severe intellectual and visual disabilities is lacking. Chapter 2 addresses the feasibility and test-retest reliability of body composition measurements of persons with SIMD.

Test–retest reliability and feasibility for most measurements of persons with SIMD are acceptable. Skinfold measurements, however, could not be reliably performed with these participants. Therefore, assessing body fat composition in adults with SIMD through skinfold measurements is not recommended. Measuring tibia length and using the determined formula to calculate body height from tibia length is a reliable alternative for measuring body height. Although measuring body height of persons with SIMD as outlined in our protocol was feasible, the feasibility of performing tibia length measurements was much better.

Furthermore, our results indicate that according to Body Mass Index (ratio between body height and body weight, BMI) values, 10% of the female participants was obese while none of the male participants was obese. When waist circumference was used as a criterion, 39% of the female and 7% of the male participants was classified as obese.

Waist circumference measured in supine position

Waist circumference as an indicator of abdominal fat is an important predictor of health risks. In healthy participants, waist circumference is measured in standing position. It is unknown whether waist circumference can be measured validly and reliably when a participant is in a supine position. This assumption however is a critical one when international standards for healthy

subjects are to be applied to persons with profound intellectual, visual, and motor disabilities. Chapter 3 deals with the validity and reliability of measuring waist circumference of persons with PIMD.

The validity study performed with healthy participants, during which we compared waist circumference obtained in standing and supine positions, revealed significant differences between standing and supine waist circumference measurements. We found that the validity of supine waist circumference is biased towards higher values (1.5 cm) of standing waist circumference. However, standing waist circumference can be predicted from supine measurements using a simple prediction equation. This equation allows the comparison of supine measurements of disabled persons with international standards of waist circumference.

The test-retest study with PIMD participants, in which we measured the waist circumference of subjects in supine position, revealed no significant differences and showed good agreement between test and retest waist circumference values. It was concluded that supine waist circumference can be reliably measured in participants with PIMD.

Functional exercise and aerobic capacity measured with over-ground tests

Cardiorespiratory fitness can be divided into functional exercise and aerobic capacity. Timed walking tests are a valuable tool for assessing these components of physical fitness. However feasibility, validity and reliability of walking tests for persons with SIMD are so far unknown. Chapter 4 seeks to address the issue of the cardiorespiratory component of physical fitness. Therefore, a study is put forward with the purpose of examining the feasibility and test-retest reliability of both the six-minute walking distance test (6MWD) as an adapted shuttle run test (aSRT) for participants with SIMD. SIMD participants performed the 6MWD and the aSRT twice while wearing a heart rate monitor.

The results show that the 6MWD is feasible and reliable for measuring functional exercise capacity of all participants with SIMD. The aSRT is feasible and reliable for measuring aerobic capacity of participants with the highest motor functioning level in the study population. In addition, we found that the participant's motivational level can influence test outcomes, so we recommend to include both heart rate monitoring and motivational score into the protocols of the aSRT and 6MWD.

Furthermore, we compared the mean distance of the 6MWD as executed by our participants with values reported in other studies. This comparison indicated that persons with SIMD performed poorer on the 6MWD than those with other specific (chronic) health conditions including heart failure or COPD. The poor 6MWD results we observed suggest that a low functional exercise capacity of persons with severe multiple disabilities is a serious health problem, which in turn can burden their independence in day-to-day activities. Based on this result, further research should be aimed at developing, implementing and evaluating an appropriate intervention aimed at reducing health problems related to low functional exercise capacity and low aerobic capacity.

Aerobic capacity measured with a graded treadmill test

Exercise tests using treadmills are a valuable tool for assessing aerobic capacity. However, a treadmill protocol for persons with SIMD is not yet available.
Chapter 5 examines the feasibility, validity and reliability of the adapted Shuttle Run Test performed on a treadmill by participants with SIMD.

Participants with SIMD performed the aSRT on a treadmill twice and a validity test was performed afterwards. Our results show that the feasibility, validity and test-retest reliability were sufficient for the aSRT on the treadmill when testing participants with SIMD. For determining peak heart rate of SIMD individuals, the validity of the aSRT on the treadmill was better than that of the aSRT performed over ground.

An equation used for estimating peak heart rate for people with ID systematically overestimates peak heart rate for people with SIMD. Therefore, it is recommended to adjust this equation in future research so as to enable a better prediction of the peak heart rate of this specific group.

Balance Scale

Sufficient balance is necessary to perform daily activities. Since the feasibility and reliability of balance tests are so far unknown for persons with SIMD, chapter 6 describes a study with the purpose of determining the feasibility and reliability of the modified Berg Balance Scale (mBBS) for this specific group.

Participants with SIMD performed the mBBS twice with a one week interval. The results show that the test-retest reliability of 10 out of 12 mBBS tasks is acceptable. The mBBS is thus a reliable test for evaluating the functional balance of persons with SIMD. Furthermore, the mBBS is a feasible instrument for the tested target group.

Tests to measure muscle tone and level of spasticity

The quality of daily movement depends partly on muscle tone or level of spasticity, which can be measured by the Modified Ashworth Scale (MAS) and the Modified Tardieu Scale (MTS). However, no research has been performed to determine the psychometric properties of the MTS and the MAS for persons with PIMD. The purpose of the study described in Chapter 7 was to determine the feasibility, the test-retest and interrater reliability of the MAS and the MTS in persons with PIMD.

Participants with PIMD were measured twice using both the MAS and the MTS, with a one-week interval between test and retest. Two observers performed the measurements.

The data indicated that the feasibility of the MAS and MTS for measuring muscle tone in persons with PIMD was good. For both test-retest and interrater reliability, measurements of the MAS revealed acceptable agreement. However, for both test-retest and interrater reliability, the measurements of the MTS showed insufficient agreement. The MAS may be a good method to evaluate the quality of daily movements of individuals with PIMD. A good instruction may contribute to a better reliability.

Heart rate patterns

Assessing physical activity levels of persons with PIMD is important, but these levels are difficult to reliably quantify when dealing with participants who are not able to walk. Heart rate monitoring

may be an indicator of activity levels, however, both the method of dating heart rate patterns as the correlation between heart rate and activity level for this specific group had so far not been subject to research.

Chapter 8 describes heart rate monitoring and heart rate patterns of persons PIMD. Furthermore, this chapter examines the relative activity of persons with profound intellectual and multiple disabilities when compared to the American College of Sports Medicine guidelines of healthy physical activity. It also looks at the correlation between heart rate patterns and level of activity of this specific target group. Finally, the influence of covariates such as gender, age, and common co-morbidity on heart rate height are examined and participants are classified according to heart rate height during physical activity.

Using a heart rate monitor, heart rate patterns were measured 8 hours during 6 days. Heart rate intensity was calculated using heart rate reserves. Physical activity levels were also measured with questionnaires filled out by the caregivers of the participants.

Our study shows that persons with PIMD are not sufficiently physically active based on the guidelines of American College of Sports Medicine. Time of day and age have considerable influence on heart rate patterns. We observed four classes in heart rate patterns of persons with PIMD.

Discussion

Chapter 9 summarizes the main findings of this research and puts them in perspective. The most important results are mentioned and discussed against the background of other research. The main products of this thesis are feasible, valid and reliable tests of physical fitness for persons with SPIMD, which can directly be implemented into daily practice. The main conclusion is that persons with SPIMD are able to learn and become accustomed to test and measurement situations if an optimal test environment is created. Various tested ways to enhance test environment are described, and suggestions for further possible enhancement of test conditions are put forward. Furthermore, this thesis has brought forth a couple of important findings regarding formulas or equations applicable to research on this specific target group.

The benefit of combining both severity of ID and GMFCS level to classify the abilities of persons with severe or profound ID is once more underscored. Moreover, the steps necessary to explore the concept of participation for persons with SPIMD as well as the relation between participation and physical fitness, activity and health are described.

Finally, methodological issues are discussed, followed by the implications for clinical practice and recommendations for further research. The main recommendation of this research is to develop, perform and evaluate tailored interventions with the aim of promoting components of physical fitness and participation for individuals with SPIMD.

Samenvatting

Lijst met afkortingen en betekenissen

ZEVMB	(Zeer) ernstige verstandelijke, visuele en motorische (meervoudige) beperkingen: term voor de hele doelgroep van dit proefschrift.
ZEVB	(Zeer) ernstige verstandelijke en visuele beperkingen.
ZEMB	(Zeer) ernstige verstandelijke, visuele en motorische (meervoudige) beperkingen: de groep mensen voor wie deze aanduiding geldt, kunnen in het algemeen niet staan en lopen.
Validiteit	Een test is valide, als deze meet wat deze beoogt te meten.
Betrouwbaarheid	Een test is betrouwbaar, 1) als twee keer kort na elkaar meten, dezelfde uitkomst geeft (test-hertestbetrouwbaarheid); en/of 2) als twee mensen dezelfde uitkomst scoren op de test (interbeoordelaarsbetrouwbaarheid).
ACSM	American College of Sports Medicine: centrum waarin begrippen rondom fysieke fitheid en bewegingsactiviteiten voor verschillende doelgroepen worden gedefinieerd.
BMI	Body Mass Index, de verhouding tussen lichaamslengte en lichaamsgewicht
6MWT	6 Minuten Wandel Test: afstand die een persoon in 6 minuten aflegt, waarmee het uithoudingsvermogen kan worden bepaald.
aSRT	aangepaste Shuttle Run Test: wandeltest met oplopende snelheid om het uithoudingsvermogen te bepalen.
mBBS	aangepaste Berg Balans Schaal: balans test.
MAS	Modified Ashworth Scale: een test om spierspanning mee te meten.
MTS	Modified Tardieu Scale: een test om spierspanning mee te meten

Inleiding

Voldoende bewegen en fitheid zijn voor iedereen van groot belang voor een goede gezondheid, welbevinden en kwaliteit van leven. Daarnaast zorgen bewegen en fysieke fitheid voor betere mogelijkheden om actief mee te kunnen doen aan dagelijkse activiteiten en situaties. Deze actieve deelname aan dagelijkse bezigheden en situaties wordt participatie genoemd.

Voor mensen met (zeer) ernstige verstandelijke en visuele beperkingen bestaat participatie vooral uit het hebben van invloed op dagelijkse bezigheden en situaties. Voorbeelden hiervan zijn het kiezen van kleding, broodbeleg, gaan zwemmen of gaan wandelen; het uitvoeren van een bepaalde deeltaak binnen een grotere taak; het zelf meehelpen met aankleden. Het is belangrijk dat deze mensen, met de beperkte mogelijkheden die ze hebben, actief mee kunnen doen aan dagelijkse bezigheden en situaties of in ieder geval invloed hebben op het verloop van hun activiteiten. Dit is echter voor hen niet vanzelfsprekend door de combinatie van de (zeer) ernstige verstandelijke en visuele beperkingen. Daarnaast hebben deze mensen een verhoogd risico op het krijgen van gezondheidproblemen, vaak ook als gevolg van beperkingen in het bewegen (motorische beperkingen).

Dit roept de volgende vragen op: *'Als fit zijn van groot belang is, hoe is het dan gesteld met de fitheid van mensen met (zeer) ernstige verstandelijke, visuele en motorische beperkingen?'* En: *'Hoe kan dit gemeten worden?' 'Als een persoon niets tot heel weinig ziet en daarnaast weinig begrijpt van zijn of haar omgeving, hoe is het dan om gemeten en getest te worden? En hoe gemotiveerd is die persoon dan om metingen te ondergaan en testen zo goed mogelijk uit te voeren?'*

Deze vragen vormden de aanleiding voor het opzetten van de onderzoeken voor dit proefschrift, met als doel het verkrijgen van inzicht in de mate van fitheid van mensen met (zeer) ernstige verstandelijke, visuele en motorische beperkingen. In eerste instantie richt dit proefschrift zich op het aanpassen van bestaande testprocedures en meetmethodes voor deze specifieke doelgroep. In de tweede plaats richt dit proefschrift zich op de uitvoerbaarheid, betrouwbaarheid en validiteit van metingen en testen voor fysieke fitheid en hoeveelheid bewegingsactiviteiten bij mensen met (zeer) ernstige verstandelijke, visuele en motorische beperkingen.

Mensen met verstandelijke beperkingen

Er zijn verschillende niveaus in de mate van verstandelijke beperkingen te onderscheiden: lichte, matige, ernstige en zeer ernstige verstandelijke beperkingen. Dit onderzoek beperkt zich tot mensen met ernstige en zeer ernstige verstandelijke beperkingen. Mensen met ernstige verstandelijke beperkingen hebben een intelligentie vergelijkbaar met die kinderen van tussen de 3 en 6 jaar en zijn daardoor aangewezen op continue begeleiding bij dagelijkse bezigheden. Mensen met zeer ernstige verstandelijke beperkingen hebben een intelligentie die vergelijkbaar is met kinderen jonger dan 3 jaar oud. Mensen met een zodanige verstandelijke beperking hebben problemen met hun eigen verzorging, zindelijkheid, communicatie en mobiliteit en zijn volledig aangewezen op hulp en begeleiding van anderen.

In het algemeen zijn bij mensen met verstandelijke beperkingen gezondheidsrisico's groter en zij hebben een grotere kans op bijkomende stoornissen als beperkingen in het gezichtsvermogen of het gehoor (visuele of auditieve beperkingen), epilepsie, neurologische en motorische beperkingen. Naarmate de ernst van de verstandelijke beperking toeneemt, komen deze bijkomende stoornissen ook vaker voor.

Uit onderzoek is gebleken dat mensen met lichte of matige verstandelijke beperkingen vaak onvoldoende bewegen om fysiek fit en gezond te blijven. De kwaliteit van bewegen is minder goed dan bij mensen zonder beperkingen. Het is zelfs zo, dat naarmate de ernst van de verstandelijke beperking toeneemt, de kwaliteit van bewegen verder afneemt. De samenwerking tussen zintuigen en het bewegen is in veel gevallen verstoord. Vaak leiden deze mensen een inactief bestaan: 50% van de mensen met een verstandelijke beperking in Europa is inactief. Mede daardoor neemt de kans op (ernstig) overgewicht toe, hetgeen onder andere hart- en vaatziekten en diabetes mellitus type II kan veroorzaken. Concluderend kunnen we stellen dat door inactiviteit en verminderde kwaliteit van bewegen, de fysieke fitheid van mensen met verstandelijke beperkingen vaak onvoldoende is.

Fysieke fitheid van mensen met (zeer) ernstige verstandelijke en visuele beperkingen

Het blijkt, dat ook mensen met visuele beperkingen vaak onvoldoende bewegen om fit te blijven en dat de kwaliteit van bewegen minder goed is dan bij mensen met een goede visus. Hierdoor hebben mensen met visuele beperkingen vaak een lagere fitheid dan mensen met een goede visus. 92% Van de mensen met (zeer) ernstige verstandelijke beperkingen heeft ook een visuele beperking. Deze mensen hebben minder compensatiemogelijkheden en de twee beperkingen versterken elkaar: daarom wordt er niet gesproken van een dubbele, maar van een meervoudige beperking. Een logisch gevolg van de voorgaande bevindingen is dat mensen met een combinatie van verstandelijke en visuele beperkingen een nog groter risico hebben op problemen met het bewegen, fitheid en participatie.

Echter, er is nog weinig bekend over de hoeveelheid bewegingsactiviteiten en de fysieke fitheid bij mensen met (zeer) ernstige verstandelijke, visuele en motorische beperkingen (ZEVMB), omdat er nog geen uitvoerbare, betrouwbare en valide testen beschikbaar zijn. Als gevolg van beperkingen in verstandelijk en sociaal functioneren, is het moeilijk om de fysieke fitheid en activiteit te meten bij deze mensen. Een belangrijke reden is dat mensen met ZEVMB allereerst niet begrijpen wat het woord 'testen' inhoudt en wat er van hen verwacht wordt. Daarnaast zijn ze niet gewend aan testen en ze begrijpen vaak de testinstructies niet. De visuele beperking versterkt dit alles nog omdat voordoen van een testopdracht daarbij niet kan helpen. Als een persoon testinstructies niet begrijpt, dan geeft de betreffende test geen goed beeld van wat iemand werkelijk kan. In dat geval meet de test niet wat deze zou moeten meten en is daardoor niet valide. De uitvoerbaarheid van fysieke fitheidtesten en – metingen heeft daarom extra aandacht nodig. Het is van groot belang dat er aandacht is voor de testomgeving en de testinstructies: is er een veilige en bekende omgeving, zijn de instructies begrijpelijk over te dragen en wie draagt ze over? Verder is een persoon met ZEVMB vaak niet gemotiveerd om optimaal te presteren, omdat hij of zij niet begrijpt wat het doel is van zo'n test of meting. Het kan helpen om deze mensen te laten oefenen voor de testen, zodat ze gewend raken aan de testprocedure.

Een andere factor die het testen en meten van mensen met ZEVMB bemoeilijkt, zijn problemen met het bewegen en vergroeiingen van de wervelkolom of andere gewrichten. Als gevolg hiervan kunnen deze mensen vaak niet voldoende of helemaal niet rechtop staan, waardoor meetprocedures aangepast moeten worden. Daarom is het van belang om voor elke test af te spreken voor welke doelgroep deze bedoeld is, bijvoorbeeld voor mensen met een bepaald

verstandelijk niveau en bepaalde motorische vaardigheden.

In samenspraak met experts uit het werkveld voor mensen met ZEVMB zijn de volgende relevante kenmerken van fysieke fitheid gedefinieerd: lichaamsverhoudingen, uithoudingsvermogen, balans, spierspanning en spierkracht.

Zoals eerder opgemerkt, zijn de motorische vaardigheden van deze doelgroep erg verschillend. Het zijn juist deze vaardigheden die ook bepalen welke testen en metingen van belang zijn voor en uitgevoerd kunnen worden door deze mensen.

Daarom wordt voor deze mensen een motoriekclassificatiesysteem gehanteerd, waarbij de mensen naar motorische mogelijkheden in categorieën worden ingedeeld.

In de studies waaraan vooral mensen met ernstige verstandelijke beperkingen deelnemen, wordt de doelgroep genoemd: mensen met zeer ernstige verstandelijke en visuele beperkingen (ZEVB). In de studies waaraan vooral mensen met zeer ernstige verstandelijke beperkingen deelnemen, wordt de doelgroep genoemd: mensen met (zeer) ernstige verstandelijke, visuele en motorische (meervoudige) beperkingen (ZEMB).

Doel en onderzoeksvragen van dit proefschrift

Het doel van dit proefschrift is het bepalen van de uitvoerbaarheid, de validiteit en de betrouwbaarheid van fitheidtesten bij mensen met ZEVB [hoofdstukken 2, 4, 5 en 6] en ZEMB [hoofdstukken 3, 7 en 8]. De volgende zeven hoofdonderzoeksvragen zijn opgesteld:

1. Zijn metingen om de lichaamsverhoudingen te bepalen bij mensen met ZEVB uitvoerbaar en betrouwbaar [hoofdstuk 2]?
2. Is de tailleomvang in liggende positie valide en betrouwbaar te meten bij mensen met ZEMB [hoofdstuk 3]?
3. Is het meten van uithoudingsvermogen met twee verschillende wandeltesten bij mensen met ZEVB uitvoerbaar en betrouwbaar [hoofdstuk 4]?
4. Is een aangepaste wandeltest op de loopband bij mensen met ZEVB uitvoerbaar, valide en betrouwbaar [hoofdstuk 5]?
5. Is een aangepaste balansschaal bij mensen met ZEVB uitvoerbaar en betrouwbaar [hoofdstuk 6]?
6. Zijn twee testen om de spierspanning te meten uitvoerbaar en betrouwbaar bij mensen met ZEMB [hoofdstuk 7]?
7. Zijn mensen met ZEMB voldoende fysiek actief op basis van hartslagpatronen vergeleken met de norm van de American College of Sports Medicine; kan de mate van fysieke activiteit betrouwbaar bepaald worden met hartslagmeting; zijn de hartslagpatronen te classificeren en beïnvloedende factoren op de hartslaghoogte te vinden [hoofdstuk 8]?

Metingen van lichaamsverhoudingen

Metingen van de lichaamsverhoudingen worden algemeen gebruikt om bij gezonde mensen de lichaamsverhoudingen en daarmee gerelateerde gezondheidsrisico's te bepalen. Echter, het is niet bekend of deze metingen bij mensen met ZEVB uitvoerbaar en betrouwbaar zijn. Daarom zijn in het tweede hoofdstuk de uitvoerbaarheid en test-hertest betrouwbaarheid onderzocht van verschillende metingen om lichaamsverhoudingen te bepalen bij mensen met ZEVB.

Het blijkt dat de uitvoerbaarheid en de test-hertest betrouwbaarheid voor de meeste metingen acceptabel waren bij mensen met ZEVB. Echter,

het wordt afgeraden om vetpercentage bij mensen met ZEVB te meten met de huidplooimethode omdat deze meting niet betrouwbaar was bij de deelnemers aan het onderzoek. Het meten van de onderbeenlengte en het daarmee berekenen van de lichaamslengte, bleek een veel beter uitvoerbaar en betrouwbaar alternatief voor het op de reguliere wijze meten van lichaamslengte.

Verder is uit de resultaten gebleken dat de verhouding tussen lichaamslengte en lichaamsgewicht, de Body Mass Index (BMI) bij 10% van de vrouwelijke deelnemers duidt op obesitas, terwijl geen van de mannen obesitas heeft. Echter, gemeten met de tailleomvang, blijkt dat 39% van de vrouwelijke en 7% van de mannelijke deelnemers obesitas heeft.

Liggend gemeten tailleomvang

Tailleomvang is een meting die gebruikt wordt om gezondheidsrisico's ten gevolge van overgewicht in kaart te brengen. Met de tailleomvang kan een inschatting van de hoeveelheid buikvet worden gemaakt. Een te hoge tailleomvang betekent een grotere kans op bijvoorbeeld hart- en vaatziekten en diabetes mellitus type II. Bij mensen zonder beperkingen wordt tailleomvang gemeten in staande positie. Echter, als mensen als gevolg van motorische beperkingen niet in staat zijn om te staan, kan de meting in rugligging uitgevoerd worden. De vraag is of deze meting valide is en betrouwbaar bij mensen met ZEMB. Deze vraag is relevant omdat alleen dan de liggend gemeten tailleomvang vergeleken kan worden met internationaal vastgestelde normen. Hoofdstuk 3 beschrijft de studie naar validiteit- en betrouwbaarheid van de liggend gemeten tailleomvang.

Om de validiteit van de liggend gemeten tailleomvang te bepalen, zijn bij 160 gezonde deelnemers staand en liggend verkregen metingen vergeleken met elkaar.

Het bleek dat tussen deze twee metingen duidelijke verschillen bestonden, waarbij de liggend gemeten tailleomvang gemiddeld 1,5 cm kleiner was. Met een statistisch bepaalde formule bleek het echter mogelijk om te voorspellen uit de liggend gemeten tailleomvang, wat de staand gemeten tailleomvang zou zijn geweest. Hierdoor is het mogelijk om de internationaal geldende normen toe te passen bij mensen met ZEMB.

In een test-hertest studie zijn daarna bij mensen met ZEMB twee metingen die in liggende positie verkregen waren met een week tussenpauze, met elkaar vergeleken. Het blijkt dat deze metingen betrouwbaar zijn bij deze doelgroep.

Inspanningsvermogen getest met een veldtest

Bij mensen zonder beperkingen worden een 6 Minuten Wandel Test (6MWT) of een Shuttle Run Test vaak gebruikt om inspanningsvermogen in kaart te brengen. Het is echter niet bekend of de uitvoerbaarheid en betrouwbaarheid van deze testen bij mensen met ZEVB voldoende is. Hoofdstuk 4 beschrijft de test-hertest studie van de 6MWT en een aangepaste Shuttle Run Test (aSRT) bij mensen met ZEVB. Mensen met ZEVB hebben de 6MWT en de aSRT twee keer, met een week tussenpauze, uitgevoerd. Alle deelnemers droegen een hartslagmeter.

De 6MWT bleek uitvoerbaar en betrouwbaar bij alle mensen met ZEVB. Echter, bij de groep mensen die minder goed kan lopen, bleek de aSRT niet betrouwbaar te zijn.

De deelnemers aan het onderzoek hebben nog niet optimaal gepresteerd. Dit blijkt uit het feit dat zij niet hun maximale hartslagen behaalden tijdens de testen. De begeleiders, die de deelnemers goed kennen, hebben ook ingeschat hoe de motivatie van de deelnemers tijdens de testen was.

Hieruit bleek, dat de motivatie van de deelnemers de testuitkomsten beïnvloedt en daarom is het van belang de motivatiescores in beide testprotocollen te integreren.

Daarnaast zijn de gemiddeld behaalde afstand van de deelnemers tijdens de 6MWT vergeleken met de afstanden behaald door andere doelgroepen. Deze vergelijking wijst uit dat mensen met ZEVB veel minder goed presteerden dan mensen met andere chronische aandoeningen, bijvoorbeeld mensen met longproblemen of hartfalen. Onvoldoende inspanningsvermogen zou tot ernstige gezondheidsproblemen kunnen leiden, waardoor participatie aan dagelijkse bezigheden zou kunnen afnemen. Gebaseerd op deze bevindingen is het van belang om passende beweegprogramma's gericht op het verbeteren van inspanningsvermogen voor mensen met ZEVB te ontwikkelen en te evalueren.

Inspanningsvermogen getest op de loopband

Andere doelgroepen behalen tijdens een Shuttle Run Test uitgevoerd op de loopband hun maximale hartslag en daardoor ontstaat een valide beeld van hun inspanningsvermogen. Tijdens de aSRT in de gymzaal bleek dat mensen met ZEVB niet hun maximale hartslag behaalden. De vraag is nu of dit wel lukt op de loopband, waardoor de validiteit van deze test toeneemt. De uitvoerbaarheid, validiteit en betrouwbaarheid van de aSRT op de loopband waren nog niet geëvalueerd bij mensen met ZEVB. Het onderzoek naar deze aspecten staat beschreven in hoofdstuk 5. De aSRT is twee keer met een week tussenpauze uitgevoerd door mensen met ZEVB om de uitvoerbaarheid en de test-hertest betrouwbaarheid van de aSRT te kunnen bepalen. Daarnaast is aan de hand van de hoogte van de maximale hartslag gekeken of de aSRT op de loopband ook valide is om uithoudingsvermogen van mensen met ZEVB te bepalen.

De uitvoerbaarheid en de test-hertest betrouwbaarheid van de aSRT bleken voldoende. De validiteit van de aSRT op de loopband bleek beter te zijn dan van de aSRT in de gymzaal. Voor het berekenen van de maximale hartslag wordt bij mensen zonder beperkingen vaak de formule: '220 –leeftijd' gehanteerd. Voor mensen met een verstandelijke beperking is een aangepaste formule berekend in eerder onderzoek. Echter, deze formule blijkt systematisch de maximale hartslag van mensen met ZEVB te overschatten. Daarom wordt aanbevolen om deze formule in een volgend onderzoek aan te passen, zodat er een betere voorspelling van de maximaal te behalen hartslag bij mensen met ZEVB te maken is.

Balanstest

Voldoende evenwicht, of balans, is voorwaarde voor het uitvoeren van dagelijkse bezigheden. De Berg Balans Test evalueert de balans bij verschillende doelgroepen. Echter, de uitvoerbaarheid en betrouwbaarheid van een balanstest zijn nog niet geëvalueerd bij mensen met ZEVB. Hoofdstuk 6 beschrijft het onderzoek hiervan bij een aangepaste Berg Balans Schaal (mBBS). Deze test is twee keer met een week tussenpauze uitgevoerd door mensen met ZEVB.

De uitvoerbaarheid en de test-hertest betrouwbaarheid van 10 van de 12 taken van de mBBS bleken voldoende bij mensen met ZEVB. De 2 taken die niet betrouwbaar waren, zijn uit het protocol van de mBBS gehaald.

Spierspanningstesten

De hoogte van de spierspanning bepaalt hoe gemakkelijk een persoon met ZEMB meehelpt met bijvoorbeeld de dagelijkse bezigheden als douchen en aankleden. De hoogte van de spierspanning kan gemeten worden met de Modified Ashworth Scale (MAS) en de Modified Tardieu Scale (MTS). Echter, de uitvoerbaarheid en betrouwbaarheid van de MAS en de MTS zijn nog niet geëvalueerd bij mensen met ZEMB. Hoofdstuk 7 beschrijft het onderzoek hiernaar.

De metingen van de MAS en de MTS zijn bij 35 mensen met ZEMB twee keer met een week tussenpauze door twee beoordelaars uitgevoerd. Van de MAS bleken de uitvoerbaarheid en de betrouwbaarheid voldoende, echter van de MTS onvoldoende bij mensen met ZEMB. De conclusie is dat de MAS een goede methode is om de spierspanning te beoordelen. Een goede instructie van beoordelaars bevordert de betrouwbaarheid.

Metingen van hartslagpatronen

Het bepalen van de hoeveelheid bewegingsactiviteiten van mensen met ZEMB is van belang. Het gebruik van een stappenteller hiervoor is niet mogelijk omdat deze mensen niet kunnen lopen. Hierdoor is de hoeveelheid bewegingsactiviteiten moeilijk vast te stellen bij mensen met ZEMB. Hartslagmeting zou een aanwijzing kunnen geven van de hoeveelheid bewegingsactiviteiten bij deze mensen. Hoofdstuk 8 beschrijft de hartslagmeting en hartslagpatronen bij mensen met ZEMB. Daarnaast wordt op basis van deze hartslagmeting de hoeveelheid bewegingsactiviteiten beschreven en vergeleken met de norm voor voldoende bewegingsactiviteiten van de American College of Sports Medicine. Tot slot beschrijft dit hoofdstuk een classificatie van hartslagpatronen en beïnvloedende factoren op hartslagpatronen.

De hartslagpatronen werden per kwartier gemeten bij 24 mensen met ZEMB, gedurende zes dagen en 8 uren per dag. Het ging daarbij om vijf doordeweekse dagen en een weekenddag. Gelijk met de hartslagmetingen, werd door woon- en activiteitenbegeleiders in een dagboekje bijgehouden hoe fysiek actief de deelnemers aan het onderzoek waren.

Hartslagmeting blijkt wel een betrouwbaar meetinstrument te zijn om bewegingsactiviteiten te meten bij mensen met ZEMB. De resultaten van het onderzoek laten verder zien dat de deelnemers aan het onderzoek niet voldoende bewegingsactiviteiten hebben, vergeleken met de norm. Er zijn 4 categorieën op basis van hartslagpatronen voor mensen met ZEMB. De indruk bestaat dat dit door de invloed van emoties en persoonlijke factoren veroorzaakt wordt. Daarnaast blijken de tijd op de dag en leeftijd een duidelijke invloed op hartslagpatronen te hebben, evenals bij mensen zonder beperkingen.

Discussie

In hoofdstuk 9 wordt een algemene discussie gevoerd over de inhoud van dit proefschrift. De resultaten van het onderzoeksproject worden nog eens kort genoemd en bediscussieerd in het licht van ander onderzoek.

De belangrijkste resultaten van dit proefschrift zijn uitvoerbare, betrouwbare en valide testen om de fysieke fitheid bij mensen met ZEVB en ZEMB vast te stellen, die direct toepasbaar zijn in de dagelijkse praktijk van de zorg voor deze mensen. Er zijn twee nieuwe formules berekend, die te gebruiken zijn bij onderzoek van deze bijzondere doelgroep.

Verder is duidelijk geworden dat mensen met ZEVB en ZEMB aan testen kunnen wennen, als de omgeving maar de juiste voorwaarden creëert. Deze juiste voorwaarden worden beschreven in de discussie. Voorbeelden hiervan zijn:

- aanpassingen aan bestaande testprotocollen;
- het uitvoeren van oefensessies om vertrouwd te raken met dit protocol;
- testen op een logische tijd, bijvoorbeeld tijdens een gymuurtje;
- vertrouwde en geschoolde begeleiding, bijvoorbeeld bewegingsagogen.

De motorische vaardigheden van mensen met mensen met ZEVMB bepalen welke testen en metingen van belang zijn voor een persoon. In de discussie komt aan de orde dat het van belang is dat naast de indeling naar verstandelijke beperking, ook een indeling naar bewegingsmogelijkheden gehanteerd wordt bij mensen met ZEVMB.

Het verband tussen fysieke fitheid, bewegingsactiviteiten, gezondheid en participatie wordt nader toegelicht. Daarnaast worden statistische aspecten bij onderzoek bij mensen met ZEVB en ZEMB nader belicht. Tot slot worden aanbevelingen gedaan voor verder onderzoek bij mensen met ZEVB en ZEMB. Hierbij komt met name de aanbeveling voor het ontwikkelen van beweegprogramma's speciaal voor mensen met ZEVB en ZEMB nadrukkelijk naar voren.

Dankwoord

Bij het opzetten en uitvoeren van toegepast wetenschappelijk onderzoek is samenwerken en communiceren van groot belang. Toen ik begon aan de onderzoeken voor dit proefschrift, heb ik ontdekt dat dit voor mij een van de leuke aspecten van onderzoek is.

Aan het eind van dit proefschrift wil ik dan ook graag alle mensen die op wat voor manier dan ook betrokken zijn geweest bij de onderzoeken, bedanken.

Allereerst bedank ik mijn begeleiders, promotores professor C.P. van der Schans en professor B. Steenbergen en copromotor dr. R. van Wijck.

Beste Cees, vanaf het eerste moment van onze kennismaking in 2005 viel mij op hoe je mogelijkheden en kansen ziet voor onderzoek, en hoe je onderzoek bereikbaar maakt voor de directe praktijk. Je vaardigheden om samenwerken te stimuleren en verbindingen te leggen tussen onderzoek en praktijk, hebben de basis gelegd voor dit toegepaste wetenschappelijke onderzoek. Ik wil je dan ook bedanken voor de kansen die je me gegeven hebt!
Ik waardeer het vertrouwen dat je in mij en mijn vaardigheden stelde. Ik vind het juist daarom ook zo mooi dat je sinds september 2010 hoogleraar en daarmee mijn eerste promotor bent. Ik heb ontzettend veel geleerd van je kennis over opzet van onderzoek, je statistische kennis en je altijd positieve en opbouwende feedback op mijn artikelen. Het is heel prettig dat je altijd snel reageert en feedback geeft. Ik ben blij dat we in de toekomst blijven samenwerken en hoop nog veel van je te leren!

Beste Bert, hoewel je pas vanaf 2009 bij ons onderzoek betrokken bent, heb ik heel veel van je duidelijke aanwijzingen geleerd. Ik bedank je voor je *open mind* voor samenwerken en de kansen die je mij met deze samenwerking gaf. Ik ben heel blij met jou als tweede promotor en je Nijmeegse en toch ook een beetje Groningse input in het onderzoek. Ik heb heel veel bewondering voor jouw kennis en vaardigheden in het schrijven van subsidievoorstellen: van het samen schrijven van het subsidievoorstel voor Inzicht heb ik veel geleerd! Maar ook de artikelen waarin je hebt meegeschreven, zijn verbeterd door jouw input. Ik ben blij met je inbreng in de *General Discussion*, het is zo echt beter dan het eerst was!

Beste Ruud, jij bent mijn grote inspirator voor het onderwerp van de onderzoeken vanaf het eerste begin. Je lezing op de NGBZ dag voor fysiotherapeuten over jouw onderzoek bij mensen met lichte en matige verstandelijke beperkingen op het gebied van fysieke fitheid, heeft mijn belangstelling gewekt en nieuwsgierigheid geprikkeld. De samenwerking tussen de Hanzehogeschool en Bewegingswetenschappen binnen dit onderzoek is heel waardevol en vruchtbaar gebleken. Ik heb veel van je geleerd over de inhoudelijke kant van fysieke fitheid, je statistische kennis, je kennis van de literatuur en de doelgroep. Je waardevolle opmerkingen over de artikelen en de tips voor relevante tijdschriften zijn van onschatbare waarde gebleken. Ik hoop dat ik nog gebruik mag blijven maken van je expertise! Bijzonder was de rol die Fleur, jouw dochter, aan het eind van het promotietraject nog speelde. Zij heeft niet alleen verschillende stukken gecorrigeerd op Engels, maar ook samen met jou en mij meegedacht over de inhoud, wat een grote meerwaarde bleek!

De leden van de leescommissie, professor H. M. Evenhuis, professor C. Vlaskamp en professor G. Lancioni dank ik hartelijk voor het beoordelen van het proefschrift, helemaal geweldig dat het zo snel kon!

Koninklijke Visio De Brink, haar cliënten en cliëntvertegenwoordigers en al haar medewerkers ben ik veel dank verschuldigd voor alle kansen die mij geboden zijn. Ik ben trots op De Brink, omdat ze dit onderzoek heeft aangedurfd en mogelijk gemaakt om daarmee extra expertise over onze bijzondere doelgroep te vergaren.

Allereerst wil ik dan ook onze cliënten heel hartelijk bedanken voor het meedoen aan het onderzoek en de grote groep cliëntvertegenwoordigers voor het toestemming geven hiervoor. Ik vond het vertrouwen dat men hiermee in mij stelde hartverwarmend! Ik geniet nog iedere dag van het werken met onze bijzondere doelgroep!

Gini Verheggen, eerst als collega fysiotherapie, later als manager, heb jij dit onderzoek altijd ondersteund en toegejuicht: 'Ga dit project maar trekken', was toen je advies waar ik nu nog heel blij mee ben. Loes Hogenhuis, manager van het Expertisecentrum van De Brink, ik heb enorm veel bewondering voor de wijze waarop jij het expertisecentrum van De Brink hebt vormgegeven, opgebouwd en uitgebouwd. Je bent ontzettend goed in het leggen van verbindingen en in het faciliteren van werkzaamheden van professionals. Samen met Ymkje de Vries, coördinerend begeleider van het Expertisecentrum, vormde je een super goed team. Ik heb met heel veel plezier met jullie samen gewerkt. Bedankt voor alle kansen en het plaveien van de weg!

Sanny van der Steen, mijn begeleider tijdens de laatste twee jaren van het onderzoek, jouw begeleiding vormde een heel waardevolle aanvulling, van je goede ideeën en het inhoudelijk meedenken heb ik veel geleerd! Dank je wel! Ik ben je ook heel dankbaar voor je inzet voor mijn vervolgtraject bij Koninklijke Visio. Voor de manier waarop je dit hebt voorbereid heb ik grote bewondering. Ook KEI wil ik bedanken voor het waarborgen van de voortgang van het onderzoek de afgelopen twee jaren.

Jo Triepels, eerst in de Raad van Bestuur van De Brink, nu van Koninklijke Visio, ik dank jou allereerst heel hartelijk voor je inzet voor dit onderzoek en je inhoudelijke interesse, en ten tweede voor je visie op de toekomst van onderzoeksmogelijkheden binnen Koninklijke Visio en de mogelijkheden die je mij hiermee biedt.

Lex van Hemert, directeur van De Brink, ik bedank jou voor de wijze waarop je het onderzoek hebt ondersteund en ook nu nog het vervolg ondersteunt en faciliteert.

De managers van de woningen en dagbesteding wil ik bedanken voor alle support en zinvolle tips bij het uitvoeren van de onderzoeken. De medewerkers Wonen en Dagbesteding hebben veel werk voor het onderzoek verzet in de vorm van invullen van allerlei vragenlijsten, feed back geven, meedenken met plannen van onderzoeken en niet te vergeten het begeleiden en geruststellen van de cliënten, bedankt hiervoor! En jullie zijn nog niet van me af!

Voor de collega´s van beweging, Janine Wolting, Eric van Gelderen, Henny Honebeek, Jackie van Dalfsen, Marjan Gritter, Marije Warners en vervangers heb ik tijdens het onderzoek nog meer bewondering gekregen dan ik al had: dankzij de vanzelfsprekende en uitlokkende begeleiding van alle cliënten tijdens toch vaak spannende testen zijn de onderzoeken succesvol en uitvoerbaar gebleken. Ik weet wel dat jullie af en toe zuchtten als ik weer aankwam met een onderzoek, maar jullie hebben het steeds klaar gespeeld om ieder onderzoek tot een succes te maken. Bedankt! Jullie kunnen trots zijn op jullie expertise en het is goed dat jullie dit blijven uitdragen!

En dan mijn collega's van fysiotherapie, Gineke Hanzen, Peter Dijkhuis en Jetty Mensies, heel heel hartelijk dank voor deze kans die jullie mij zo van harte gunden en ook voor alle ondersteuning, support, de luisterende oren en humor. Ik tref het maar enorm met zulke collega's. Jullie hebben meegedacht en meegewerkt met het onderzoek, studenten ingewerkt, heel veel namen van studenten getracht te onthouden en het kantoor gedeeld met ze. Maar zeker ook hebben jullie bijgedragen aan de inhoud en de vorm van de onderzoeken. Ik had jullie graag alle drie als paranimf gekozen.

Theo Kiestra, wat heb ik veel van jou geleerd, ik herinner me nog de eerste gesprekken over de cliënten in jouw kamer in de Boerderij toen ik net begon bij De Brink. Van jou heb ik de betekenis van een meervoudige beperking geleerd. Daarnaast heb je een waardevolle bijdrage geleverd aan de eerste artikelen op het gebied van statistiek. Heel hartelijk dank voor alles!

De collega's van orthopedagogiek wil ik heel hartelijk danken voor de inhoudelijke support en het steeds weer kijken naar lijsten met cliënten om contra-indicaties te bepalen. Ik wil graag met name Mariska van Panhuis noemen, die vanaf het begin betrokken is geweest bij de onderzoeken en adviseerde om de motoriek van de cliënten naar GMFCS niveaus in te delen.

Ook de Artsen Verstandelijk Gehandicapten, Jeroen Auener en Marijke Nieberg hebben erg bijgedragen aan de inhoud van de onderzoeken. De medische contra-indicaties zijn zo heel goed verwoord en staan in alle artikelen vermeld.

Jeromia Kramer, diëtiste, heeft meegedacht over en meegeholpen met alle onderzoeken op het gebied van lichaamssamenstelling. Jeromia, het was heel leerzaam om jouw relativerende visie op de onderzoeken te horen, dank je wel!

Herman Bezema, jou wil ik van harte bedanken voor het aanvragen en verkrijgen van de twee subsidies die de eerste twee jaren van dit onderzoek hebben mogelijk gemaakt. Joke ter Maat, Henny Kannegieter, Janny de Vries en Janny Merema, ik wil jullie allen heel hartelijk danken voor alle ondersteuning. De samenwerking met jou, Marjan van Slageren, heb ik als heel bijzonder en inspirerend ervaren, dank hiervoor!

Het lectoraat van de Hanzehogeschool wil ik bedanken voor de uren die ik in het onderzoek mocht steken. De contacten met de collega´s van de onderzoeksgroep zijn heel waardevol en leerzaam. Er heerst een prettige werksfeer in de onderzoekskamer en ik vind het altijd gezellig en inspirerend om hier te komen werken. Het is heerlijk om van gedachten te wisselen met en vragen te kunnen stellen aan mensen die zoveel kennis van zaken hebben en hetzelfde proces doormaken.

Judith van der Boom, dank je wel voor al het ondersteunen, het maken van afspraken en het meedenken met de verschillende regelzaken. Het is heerlijk om zo'n rustgevende vraagbaak te hebben en ik vond het vooral ook gezellig om af en toe even aan te wippen en een praatje te maken!

Ik ben veel dank verschuldigd aan alle studenten van verschillende opleidingen van de Hanzehogeschool, fysiotherapie, voeding- en diëtetiek en sportstudies, die een bijdrage aan het onderzoek hebben geleverd. Alleen met jullie hulp was het mogelijk om zoveel werk in zo korte tijd te verzetten! Een paar voorbeelden: het maken van planningen, meedenken over protocollen van de testen, het uitvoeren van testen en metingen in samenwerking met de afdeling beweging, invoeren van data in SPSS. Irma Evenhuis, jij hebt als eerste studente een wel heel wezenlijke bijdrage aan het begin van het project geleverd.

De samenwerking met de faculteiten bewegingswetenschappen en orthopedagogiek van de Rijksuniversiteit heeft een bijdrage aan de wetenschappelijke kwaliteit van het onderzoek geleverd, ik ben heel blij met de hulp van de studenten van deze faculteiten, hartelijk dank! Wim Krijnen, jou wil ik heel hartelijk danken voor je vakkundige hulp bij de statistiek van twee artikelen. Geweldig hoe je steeds zo snel reageerde, maar ook het antwoord voor de reviewer zo goed verwoordde! Roy Stewart, van jou heb ik veel geleerd over multilevel analysis, het was mooi dat je zoveel geduld met mij had, bedankt! Daarnaast heb je ook nog waardevolle analyses voor me uitgevoerd. Het was altijd heel gezellig om met je te bellen!

Familie en vrienden, jullie zijn zo belangrijk voor mij, dat heb ik de afgelopen jaren goed gemerkt. Het is heerlijk om successen en kleine tegenslagen te kunnen delen. Heel erg bedankt voor even een luisterend oor, maar ook voor alle afleiding, die minstens net zo belangrijk is. Een korte opsomming: Leuke familiebijeenkomsten met Kerst. Met Marga en Theresa weekendjes weg, naar de pizzeria en samen kopjes koffie drinken en vooral veel kletsen! Janny, altijd even gezellig kletsen bij een kop thee of een wijntje, bijzonder dat jij het proefschrift zo mooi hebt gemaakt. Akke, jij bent altijd zo positief en enthousiast, leuk om van alles uit te wisselen over onze kinderen en ons werk. Mirjam, heerlijk om tijdens het lopen te kletsen en vooral te lachen en zo het hoofd weer leeg te maken!

Papa en mama Heiligers, geweldig en hartverwarmend hoe jullie me in jullie gezin hebben opgenomen. Jullie zijn heel bijzonder! Wat missen we je allemaal mama en wat ongelofelijk jammer dat je er niet meer bij kunt zijn.

Papa en mama Waninge, wat ben ik ontzettend blij met jullie. Jullie onvoorwaardelijke liefde en zorg vormen een heel belangrijke basis! Het is heel geweldig hoe jullie altijd voor ons klaarstaan en ik ben ook heel erg trots op jullie!

Jelmer, Robbin, Demi en Kuki: ik ben ontzettend blij met en trots op jullie! Ik weet dat jullie soms niet snappen waar ik toch allemaal mee bezig ben. Sorry als ik weer eens met mijn hoofd niet helemaal bij jullie was, evengoed was ik met mijn hart er altijd wel!

En dan Rudy, liefste, hoewel je niet in het boekje genoemd wilde worden, hoor je er natuurlijk als geen ander in.
Jij hebt vanaf het begin al mijn verhalen aan moeten horen en had steeds een luisterend oor voor de hobbels die op het pad kwamen. Heerlijk om met jou te kunnen praten over onze cliënten. Je zegt steeds dat je niets hebt bijgedragen, maar door mijn gesprekken met jou en je kennis van onze doelgroep, kwamen de ideeën over de uitvoering van de onderzoeken. Ook je visie op lichamelijk actief zijn van de cliënten, die je steeds dagelijks in praktijk brengt, en je creativiteit daarin, vormen een voorbeeld: mogelijkheden zien in plaats van beperkingen, dat heb ik van jou geleerd. Ik hoop op nog vele gesprekken, waarin we het over participatie van onze cliënten hebben!

Curriculum vitae

Aly Waninge werd op 25 juni 1964 geboren in Zuidlaren. Na afronding van het VWO aan de Dr. Aletta Jacobsscholengemeenschap in Hoogezand in 1982, studeerde zij fysiotherapie aan de Academie voor Fysiotherapie, nu Hanzehogeschool Groningen. Na het afronden hiervan in 1986, ging zij als fysiotherapeut deeltijd werken bij De Brink in Vries. Deze deeltijd baan combineerde zij tot 1999 met het werken in een particuliere praktijk in Zuidlaren. In 1989 heeft ze de Neuro Developmental Treatment cursus voor Kinderen gevolgd, in 1992 heeft ze de opleiding Manuele Therapie aan de SOMT in Amersfoort afgerond en daarnaast heeft ze verschillende verdiepingscursussen op het gebied van fysiotherapie gedaan.
Van 2007 tot 2011 had Aly een deeltijdaanstelling als onderzoeker bij De Brink en bij het Lectoraat Transparante Zorgverlening van de Hanzehogeschool Groningen. Haar promotieonderzoek voerde ze uit onder begeleiding van prof. dr. C.P. van der Schans, lector Transparante Zorgverlening Hanzehogeschool Groningen en hoogleraar Revalidatiegeneeskunde UMCG, prof. dr. B. Steenbergen, hoogleraar Perception and Action Problems en dr. R. van Wijck, orthopedagoog, universitair docent en senior-onderzoeker Centrum voor Bewegingswetenschappen UMCG.
Het is de bedoeling dat Aly na haar promotie actief zal blijven met toegepast wetenschappelijk onderzoek binnen Koninklijke Visio, waarbij de samenwerking met de lectoraten Transparante Zorgverlening en Rehabilitatie van de Hanzehogeschool wordt geïntensiveerd.